

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 10, 2004, 11:51:21 ; Search time 5303 Seconds

(without alignments)
3210.313 Million cell updates/sec

Title: US-10-763-972-2

Perfect score: 1936

Sequence: 1 MSLILPSGRSGSRGAL.....QNKLGHPAGKRCPCGLNRS 360

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q/cgn2_1/USPRO_spool/US10763972/runat_04112004_093313_8784/app query.fasta_1.519
-DB=GenEmbl -QWMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10763972.CGN 1.1 3731 @runat_04112004_093313_8784 -NCPU=6 -ICPU=3
-NO WMAP -LARGESOURCY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_lm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1936	100.0	1082	6	BD187416 Novel pol
2	1936	100.0	1082	6	AX458306 Sequence
3	1936	100.0	1145	9	AJ345013 Homo sapi
4	1936	100.0	1334	6	AX646427 Sequence

5	1936	100.0	1334	9	AB065670 Homo sapi
6	1936	100.0	3143	6	AX365511 Sequence
7	1936	100.0	164502	2	AC026960 Homo sapi
8	1936	100.0	167084	2	AC092999 Homo sapi
9	1936	99.3	163958	2	AC021773 Homo sapi
10	1896	97.9	1108	6	AX921825 Sequence
11	1806	93.3	1019	9	AF411108 Homo sapi
12	1798	92.9	1002	6	BD095704 Novel gua
13	1785	92.2	1020	6	BD187417 Novel pol
14	1785	92.2	1020	6	AX458308 Sequence
15	1785	92.2	1076	6	AX375230 Sequence
16	1785	92.2	1076	6	AX365514 Sequence
17	1456	75.2	1140	6	AX365514 Sequence
18	1154.5	59.6	184377	10	AC121579 CQ734310 Sequence
19	1154.5	59.6	195290	2	AC135635 AC120742 Rattus no
20	1130	58.4	250610	2	AC120742 Rattus no
21	1130	58.4	276127	2	AC116265 Rattus no
22	921	47.6	510	6	AX147768 Sequence
23	921	47.6	510	6	AX521817 Sequence
24	732	37.8	681	6	AX244718 Sequence
25	681	35.2	585	6	AX244727 Sequence
26	501	25.9	1074	5	AF069555 Meleagris
27	501	25.9	1977	5	AF031897 Meleagris
28	495	25.6	1163	5	AF031897 Meleagris
29	460.5	23.8	1086	10	MMU277752 Mus muscu
30	460.5	23.8	214997	10	AL671299 Mouse DNA
31	460.5	23.8	215410	10	AC091784 Genomic S
32	453	23.4	1974	6	CQ720945 Sequence
33	453	23.4	123185	2	AP000587 Homo sapi
34	449	23.2	162059	2	AP000587 Homo sapi
35	447	23.1	2681	9	BC012104 Homo sapi
36	446.5	23.1	1651	9	HSU40223 Human vridi
37	446.5	23.1	137153	5	HSU40223 Human vridi
38	446	23.0	1134	9	AL590151 Zebrafish
39	446	23.0	1177	10	AY136753 Homo sapi
40	445.5	23.0	1232	9	Y11433 R.norvegicu
41	445.5	23.0	1429	6	X96597 H.sapiens g
42	445.5	23.0	1429	6	AG2985 Sequence 1
43	445.5	23.0	1429	6	AR270505 Sequence
44	445.5	23.0	1429	9	X91852 H.sapiens p
45	445.5	23.0	178861	2	AC068231 Homo sapi

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	COMMENT
BD187416	Novel polypeptide.	BD187416	BD187416.1	GI:32997155	JP 2003009885-A/1.	Homo sapiens (human)	Homo sapiens	1 (bases 1 to 1082)	Fidock,M.D.	Novel polypeptide	Patent: JP 2003009885-A 1 14-JAN-2003; Pfizer Ltd (EP GB) only, Pfizer Inc (US JP EP except GB)
OS	homo sapiens	PN	JP 2003009885-A/1	PD	14-JAN-2003	PF	17-DEC-2001	JP 2001382707	mark david fidock	CC	0030855.1, 17-JAN-2001 GB 0101222.8 PI

FEATURES

Source	Location/Qualifiers
1..1082	/organism="Homo sapiens"
	/mol_type="genomic DNA"

361 CTGCTCTGCAAGCTGGTGCATCTCTGTCTCTATATACCTTTACGGCAGCATCTGCTG 420

```

141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db CTGACCTGCATCTCTGTGCACCACTCTTAGGTGTGTGCACCACTGTGTGTGCTGCC 480

161 TyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
Db TACCGACCCGACGAGGATGCTGCTGGGACCAACCACTGGGCCCTGTGTGCTCTC 540

181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
Db CAGCTGCTGCCACACTGGCCTTCTCCACACGAGCTACATCAATGGCCAGATGATCG 600

201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db TATGACATGACGACGACGAGAAATTGTGATCGGCTTTTGGCTACGCGATGTTCTG 660

221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db TTGCTGGCTTTCTTTCCTTCTTGGTCATTTTGGTGTGCTATTCTACTGATGTCAG 720

241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db CCTGATCAAGCCAGAGAGAACTCTCATGAGGACAGGCAACACAGCCGACGAGTCC 780

261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
Db CCGGACCATCTACTGTGTGTGTGCTCTTACCTCTGTTTGTGCTTCTCCATATCAC 840

281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
Db TCGCTCTCTTACCTACCATCTGCTTCTGCTTCTTCTCAGGACTGCCAGCTCTGTAT 900

301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnGlnLeuProGlnPro 320
Db AGCCAGGTGTGGCTTACAGATATGAGGCTCTGTGAGTGTGAGCAGCTGCCCTCAAC 960

321 SerProValLeuSerPheLeuGlyGlyLeuAsnArgValArgLeuLeuGlnLysLeuArg 340
Db AGTCCTGTACTTCTTCAAGGGGGGCAAAATAGACTCAGGCTCTCCAGAACTGAGG 1020

341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
Db CAGAACAAAGTTGGTGAGCATCCAGCTGGGAGGAAGATGCCAGGGTTGAACAGATCT 1080

RESULT 3
HSA345013 HSA345013 1145 bp DNA linear PRI 07-SEP-2003
LOCUS Homo sapiens P2Y2-like gene for nucleoside/nucleotide receptor.
DEFINITION
ACCESSION AJ345013
VERSION AJ345013.1 GI:34495183
KEYWORDS nucleoside/nucleotide receptor; P2Y2-like gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bruss, M., Bonisch, H. and Kugelgen, I.
TITLE Molecular identification and functional characterization of a new G
protein-coupled nucleoside-/nucleotide receptor
JOURNAL Unpublished
AUTHORS Bruss, M.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2001) Bruess M., Pharmacology and Toxicology,
University of Bonn, Reuterstrasse 2 b, D-53113, GERMANY
FEATURES
Location/Qualifiers
1..1145
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/tissue_type="blood"
10..1095
gene

```

```

/ gene="P2Y2-like"
10..1095
/ gene="P2Y2-like"
/ function="G protein coupled receptor"
/ codon_start=1
/ product="nucleoside/nucleotide receptor"
/ protein_id="CAC87811.1"
/ db_xref="GI:34495184"
/ translation="MLSILLPSRSGSRGALLLEGASDEMEKMDNTSDEGLQCFSEKIKQYLSLAYSIIFILGLPUNGIVLMHSWQCTKWSCATTYLVNLMVADLLYL
LPFIITYSLDRWPFGLCKLVHFIYINLSIGLILLLTICISVGLGVGCHPLCSLP
YRTRHAWLGTSTWALVQLLEPTLAFSDTYINGQMIWYDMTQSFNFDLFAFGIV
LTLGSLFLSHLGHFGLVTDGQEPDQARGEPHEDRQSPQVHPDPTGWPLHPLFCA
LPVHSLILPHLLSAFSLGALPDQSCQGLQDMEASGECEQLPQSPVLSFKGKNRVR
LLQKLQNKLGEHPAGKRCRCLNRS"
Alignment Scores:
Pred. No.: 6,166-166 Length: 1145
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-763-972-2 (1-360) x HSA345013 (1-1145)
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
Db 10 ATGCTGTGCATTTTCTTCTTCCAGGGAGACAGAGGGAGCCGCTGTGGAGCTCTG 69
QY 21 LeuLeuGluGlyValAspArgAspMetGluLysValAspMetAsnThrSerGlnGlu 40
Db 70 CTCCTGGAGGAGCTCCCGGACATGGAGAAGTGGACATGAAATACATACAGGAACA 129
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db 130 GGTCTCTGCAGTCTTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 189
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
Db 190 ATCTTTATCTTAGGCTGCCACTAATGGCACTGCTTGTGGCACTCTCTGGGGCAACCC 249
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 250 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGGTGGCGACCTGTTAT 309
QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120
Db 310 GTGCTATTGCCCTTCTCTATCATCATCCTACTCCTAGATGACAGGTGGCCCTTCGGGAG 369
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 370 CTGCTCTGCAAGCTGTGTGCATTTCTGTGTTCTATATCAACCTTTACGGCAGCATCCTGCTG 429
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 430 CTGACCTGCATCTCTGTGCACCACTTCTAGTGTGTGCCACCACTGGGCCCTGTGTGCTCTC 489
QY 161 TyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
Db 490 TACCGACCCGACGAGGATGCTGCTGGGACCAACCACTGGGCCCTGTGTGCTCTC 549
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
Db 550 CAGCTGCTGCCACACTGGCCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTG 609
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 610 TATGACATGACGACGACGAGAAATTGTGATCGGCTTTTGGCTACGCGATGTTCTGACA 669
QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240

```

```

/ gene="P2Y2-like"
10..1095
/ gene="P2Y2-like"
/ function="G protein coupled receptor"
/ codon_start=1
/ product="nucleoside/nucleotide receptor"
/ protein_id="CAC87811.1"
/ db_xref="GI:34495184"
/ translation="MLSILLPSRSGSRGALLLEGASDEMEKMDNTSDEGLQCFSEKIKQYLSLAYSIIFILGLPUNGIVLMHSWQCTKWSCATTYLVNLMVADLLYL
LPFIITYSLDRWPFGLCKLVHFIYINLSIGLILLLTICISVGLGVGCHPLCSLP
YRTRHAWLGTSTWALVQLLEPTLAFSDTYINGQMIWYDMTQSFNFDLFAFGIV
LTLGSLFLSHLGHFGLVTDGQEPDQARGEPHEDRQSPQVHPDPTGWPLHPLFCA
LPVHSLILPHLLSAFSLGALPDQSCQGLQDMEASGECEQLPQSPVLSFKGKNRVR
LLQKLQNKLGEHPAGKRCRCLNRS"

```

CDS

ORIGIN

Db	670	TTGCTGGCTTTCTTCCCTCCTTGGTCAATTTTGGTGTGCTATTCACTGATGGTCAGGAG	729
QY	241	ProAspGlnAlaAargGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis	260
Db	730	CCTGATCAAGCCAGAGGAGAACCTCATGAGACAGGCAACACGCCCGAGGAGGTCAT	789
QY	261	ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis	280
Db	790	CGGAGCACTCTACTGCTGTGGCTCTTCACTCTGTTTGTGGCTTCCATATCAC	849
QY	281	SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly	300
Db	850	TGGCTCTCTTACCTTCACTCTGCTTCTTCTGAGACAGGCAAGGCTCTTGTGGC	909
QY	301	SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro	320
Db	910	AGCCAGTGTGGCTACAAAGATATGAGGCTCTTGGTGTGAGAGTGGCTCAACCC	969
QY	321	SerProValLeuSerPheGlyGlyGlyAsnArgValArgLeuLeuGlnLeuArg	340
Db	970	AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCCAGAAACTGAGG	1029
QY	341	GlnAsnLeuLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer	360
Db	1030	CAGAACAGTTGGTGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGACAGATCT	1089
RESULT 4			
LOCUS	AX646427	1334 bp	DNA
DEFINITION	Sequence 619 from Patent EP1270724.	linear	PAT 04-MAR-2003
ACCESSION	AX646427		
VERSION	AX646427.1	GI:28798808	
KEYWORDS	Human sapiens (human)		
SOURCE	Human sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.		
TITLE	Guanosine triphosphate-binding protein coupled receptors		
JOURNAL	Patent: EP 1270724-A 619 02-JAN-2003; National Institute of Advanced Industrial Science and Technology (JP); Center for Advanced Science and Technology Incubation, Ltd. (JP)		
FEATURES	Location/Qualifiers		
source	1..1334		
CDS	/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" join(201..715,834..1134) /note="unnamed protein product" /codon_start=1 /protein_id="CAD69342.1" /db_xref="GI:28798808" /translation="MEKVDNMTSQEQLQCFSESKYQVYLSLAYSIIFILGHLPLNGTV LMSWQGTQKRWSCATVYLVNADLVLLPFLITYSLDDRPFGELLCKLVHFLF YINLYSILLITGTSVHFGVCHPLCSLPYRFRHAWLGTSTWALVVLQPLTFLP SHTDYINGQIMWSLIKPEENLMTGNTARARSIRITLLVCGFLTLCFVPFHITRSFYL TICFLLSQDQLLMAASVAKIWRPLVSVSSCLNPVLYFLSRGAKIESSSRN"		
ORIGIN			
Alignment Scores:			
Pred. No.:	7,54e-166	Length:	1334
Score:	1936.00	Matches:	360
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-763-972-2 (1-360) x AX646427 (1-1334)			
QY	1	MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu	20

Db	117	ATGCTGTCATTTTCTTCCAGGGAAGCGGAGCGCTCGTGGAGCTCG	176
QY	21	LeuLeuGlyGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGln	40
Db	177	CTCTCTGGAGGAGCCTCCCGGACATGGGAAGGTGGACATGAATACATCACAGAACAA	236
QY	41	GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle	60
Db	237	GGTCTCTGCGAGTTCTCAGAGAGTACAGCAAGTCACTCTCTCTCCCTGGCCCTACAGTATC	296
QY	61	IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr	80
Db	297	ATCTTTATCTTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC	356
QY	81	LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr	100
Db	357	AGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCGACCTGCTTTAT	416
QY	101	ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpPropheGlyGlu	120
Db	417	GTGCTATTGCCCTTCTCATCATCACTACTCACTAGATGACAGGTGGCCCTTCGGGGAG	476
QY	121	LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu	140
Db	477	CTGCTCTGCAAGCTGGTGACCTTCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTG	536
QY	141	LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro	160
Db	537	CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTCTGTTCGTGCCCC	596
QY	161	TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu	180
Db	597	TACCGGACCCGAGGACATGCTGGTGGCCACCACTGGTGGCCCTGGTGGTCTCTC	656
QY	181	GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp	200
Db	657	CAGCTGTGCTGCCACACTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGG	716
QY	201	TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr	220
Db	717	TATGACATGACGACCAAGAGAAATTTGATCGGCTTTTGTGCTACGGCATAGTTCTGACA	776
QY	221	LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu	240
Db	777	TTGCTGCTGCTTCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	836
QY	241	ProAspGlnAlaAargGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis	260
Db	837	CCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCCGAGGACGATCCAT	896
QY	261	ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis	280
Db	897	CGGAGCACTCTACTGCTGTGGCTCTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	956
QY	281	SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly	300
Db	957	TCGCTCTTCTTACCTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGTATGGC	1016
QY	301	SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro	320
Db	1017	AGCCAGTGTGGCTTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCCCAACCC	1076
QY	321	SerProValLeuSerPheLeuGlyLysAsnArgValArgLeuLeuGlnLeuArg	340
Db	1077	AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCTCCAGAAACTGAGG	1136
QY	341	GlnAsnLeuLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer	360
Db	1137	CAGAACAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGGTTGAACAGATCT	1196
RESULT 5			
LOCUS	AB065670	1334 bp	DNA
	AB065670	linear	PRI 23-JUL-2002

Homo sapiens gene for seven transmembrane helix receptor,
 isolate:CBRC7M_233.
 AB065670
 VERSION
 AB065670.1 GI:21928614
 KEYWORDS
 Homo sapiens (human)
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S.,
 Teutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
 Genome-wide discovery and analysis of human seven transmembrane
 helix receptor genes
 Unpublished
 Suwa,M.
 2 (bases 1 to 1334)
 Direct Submission
 Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research
 Center (CBRC), National Institute of Advanced Industrial Science
 and Technology (AIST); 2-41-6 Aomi Koto-Ku, Tokyo 135-0064, Japan
 (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,
 Tel:81-3-3599-8080, Fax:81-3-3599-8081)
 This sequence is a seven transmembrane helix receptor candidate
 predicted from the whole human genome sequences using our automated
 system that contains programs of gene
 finding(GeneDecoder), sequence search, motif-domain assignment and
 transmembrane helix prediction.
 And the sequence is submitted by the collaborative project between
 [Computational Biology Research Center (CBRC), National Institute
 of Advanced Industrial Science and Technology (AIST)] and [Genome
 Science Division, Research Center for Advanced Science and
 Technology (RCAST), University of Tokyo].
 Location/Qualifiers
 1..1334
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="CBRC7M_233"
 /db_xref="taxon:9606"
 /chromosome="3"
 misc_feature
 201..1134
 /product="seven transmembrane helix receptor"
 /evidence=not_experimental
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,54e-166 Length: 1334
 Score: 1936.00 Matches: 360
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-10-763-972-2 (1-360) x AB065670 (1-1334)
 Qy 1 MetLeuSerLeuLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
 Db 117 ATGCTGTCATTTTCTCTCCAGGGAAGCAGAGAGCGGAGCGCTGCTGGAGCTG 176
 Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGlu 40
 Db 177 CTCTGGAGGAGCTCCGGGACATGGAGAGTGGACATGATACATCACAGAACAA 236
 Qy 41 GlyLeuGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerile 60
 Db 237 GGTCTCTGCCAGTCTCTCAGAGAGTACAAAGCAAGTCTACCTCTCCCTGGCTACAGTATC 296
 Qy 61 IlePheLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
 Db 297 ATCTTTATCCCTAGGCTCCCACTAAATGGCACTGCTTGCGCACTCTCTGGGCGCAACC 356
 Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuTyr 100

Db 357 AAGCGCTGGAGCTGGCCACCACCTATCTGGTGAACCTGATGGTGGCCGACCTGCTTTAT 416
 Qy 101 ValLeuLeuProPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
 Db 417 GTGCTATTGGCTTTCCTCATCATCACCTACTAGATGACAGTGGCCCTTCGGGGAG 476
 Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerileLeuLeu 140
 Db 477 CTGCTCTGCAAGCTGGTGCATCTCTGTTTATATCAACCTTTACGGCAGCATCTCTG 536
 Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 Db 537 CTGACCTGCAATCTCTGTGCACCACTTCTAGGTGTGGCCACCACCTGTGTTCCTGCC 596
 Qy 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
 Db 597 TACCGACCCGACGATGCTGCTGGGACACAGCACCTGCGGCCCTGGTGGTCTCT 656
 Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
 Db 657 CAGCTGCTGCCACACCTGCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG 716
 Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
 Db 717 TATGACATGACACGACAGAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGACA 776
 Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 Db 777 TTGCTGCTCTTCTTTCT 836
 Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
 Db 837 COTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGACACAGCCGAGCCAGGTCCAT 896
 Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
 Db 897 CCGGACCATCTACTGGTGTGGGCTCTTCAACCTCTGTTTGTGGCTTCCATATCAC 956
 Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 Db 957 TGGCTCTCTTACCTCACCATCTGCTTTCTGCTTCTCAGGACTGCCAGCTTGTATGGC 1016
 Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
 Db 1017 AGCCAGTGTGGCTACAAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAACCC 1076
 Qy 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
 Db 1077 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCCAGAACTGAGG 1136
 Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
 Db 1137 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCCGGGTTGACAGATCT 1196

RESULT 6
 AX365511
 LOCUS
 DEFINITION
 Sequence 1 from Patent WO0185764.
 AX365511
 ACCESSION
 VERSION
 AX365511.1 GI:18873666
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Ramakrishnan,S.
 Regulation of human p2y-like g protein-coupled receptor
 Patent: WO 0185764-A 1 15-NOV-2001;
 Bayer Aktiengesellschaft (DE)
 Location/Qualifiers
 1..3143
 /organism="Homo sapiens"

```

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
520. .s1374
CDS
/feature="unassigned protein product"
/codon_start=1
/protein_id="CAD24312.1"
/db_xref="GI:18873667"
/translation="MEKQVMTSOGGLCOFSEKXKVYLSLAYSIIFILGLPNTV
LWHSWGQTKRSCATYLVNLVADLLYLVPFLIITYSDRRWPFGLLECKLVHFLF
KINLYGSLLLTCSHQFGLVCHPLCSLPIYRHRHAWLGTSTWLVVLQLLPTLAF
SHTDYINGQM:WYDMMSCNPNRFLPAYIVLTSGLSLGLHFGVLVFDQGEQDQAG
EPHEDRQSPQVHPDPTGVWPLHPLFCALPYHSLLLPHLLSFAFSLFDGSGQG
LQDMEASGE"
ORIGIN
Alignment Scores:
Pred. No.: 2,35e-165 Length: 3143
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-763-972-2 (1-360) x AX365511 (1-3143)
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
Db 436 ATGCTGTCCATTGCTTCTTCAGGGGAAGCAGAGCGGGAGCCCTGCTGGAGCTCTG 495
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluIysValAspVetAsnThrSerGlnGlu 40
Db 496 CTCCTGGAGGAGGCTCCCGGACATGAGAGAGGTGGACATGAATACATCACAGGAACA 555
QY 41 GlyLeuCysGlnPheSerGluLeuValTyrrLeuSerLeuAlaTyrrSerIle 60
Db 556 GGTCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTAIC 615
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheSerTrpGlyGlnThr 80
Db 616 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCGGGCCAAACC 675
QY 81 LysArgTrpSerCysAlaThrThrTyrrLeuValAsnLeuMetValAlaAspLeuTyrr 100
Db 676 AAGCGCTGGAGCTGTGCACACCTATCTGTGGAACCTGATGTGGCCGACCTGCTTTAT 735
QY 101 ValLeuLeuProPheLeuIleIleThrTyrrSerLeuAspAspArgTrpProPheGlyGlu 120
Db 736 GTGCTATTGCCCTTCTCATCATCATCTACTACTAGATGACAGGTGGCCCTTCGGGGAG 795
QY 121 LeuLeuCysIysLeuValHisPheLeuPheTyrrIleAsnLeuTyrrGlySerIleLeu 140
Db 796 CTGCTCTGCAAGCTGGGCACTCTCTGTTCTATATCAACCTTTACGGGACATCTCTGCTG 855
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 856 CTGACCTGCATCTGTGGCACTGTTCTAGTGTGTGGCACTGTTGCTGCTGCTGCTGCTG 915
QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
Db 916 TACCGGACCGGACGATGCTGCTGGTGGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 975
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrrIleAsnGlyGlnMetIleTrp 200
Db 976 CAGCTGTGCGCCACATGCGCTTCTCCACAGCAGTATCATCATGCGCAGATGATCTGG 1035
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrrGlyIleValLeuThr 220
Db 1036 TATGACATGACAGGCAAGAGATTTTGATCGGCTTTTGCCTACGGCATAGTTCTTGACA 1095
QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db 1096 TTGCTGGCTTTCTTCTCCCTTCTTGTGTCATTTTGTGTGCTATTTCACTGATGGTCAGGAG 1155

```

```

QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db 1156 CTGTATCAGCCAGAGGAGAACCTCATGAGGACAGCAACACAGCCGAGCCAGGTCCAT 1215
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTrpHis 280
Db 1216 CGGACCATCTACTGTGTGTGGCTCTTACCCCTCTGTTTGGCTTCCATATCATCAT 1275
QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
Db 1276 TGCTCTCTTCTACCTCAGCATCTGCTTCTGCTTCTCAGGACTCCAGCTCTTGATGGC 1335
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluLeuLeuProGlnPro 320
Db 1336 AGCCAGTGTGGCTTCAAGATATGAGGCTCTGTGTGAGTGTGAGCACTGCCTCAACC 1395
QY 321 SerProValLeuSerPheIysGlyGlyIysAsnArgValArgLeuLeuGlnLysLeuArg 340
Db 1396 AGTCTGTACTTCTTTCAGGGGGGCAAAAATAGAGTCAGGCTCTCCAGAACTGAGG 1455
QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
Db 1456 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGGTTCACAGATCT 1515
RESULT 7
AC026960 164502 bp DNA linear HTG 13-APR-2000
LOCUS Homo sapiens chromosome 3 clone RP11-170K4 map 3, WORKING DRAFT
DEFINITION SEQUENCE, 15 unordered pieces.
AC026960
VERSION AC026960.2 GI:7547202
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 164502)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abram,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,K., Jones,C., Kann,L., Karatas,A.,
Klein,J., Labocque,K., Lamazares,R., Landers,T., Lechoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Navlor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Willson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 13, 2000 this sequence version replaced gi:7328839.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WfBR
Web site: http://www-seq.wi.mit.edu

```

Contact: sequence submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L8491
 Center clone name: 170_K.4
 ----- Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 156927 bases at least Q40
 Consensus quality: 160376 bases at least Q30
 Consensus quality: 161800 bases at least Q20
 Insert size: 170000; agarose-1p
 Quality coverage: 4.4 in Q20 bases; sum-of-contigs
 Quality coverage: 4.5 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 2917: contig of 2917 bp in length
* 2918: gap of 100 bp
* 3018: contig of 3293 bp in length
* 6311: gap of 100 bp
* 6411: contig of 3037 bp in length
* 9448: gap of 100 bp
* 9548: contig of 4923 bp in length
* 14471: gap of 100 bp
* 14571: contig of 5958 bp in length
* 20529: gap of 100 bp
* 20629: contig of 4039 bp in length
* 24668: gap of 100 bp
* 24768: contig of 4600 bp in length
* 29368: gap of 100 bp
* 29468: contig of 6770 bp in length
* 36238: gap of 100 bp
* 36338: contig of 10040 bp in length
* 46378: gap of 100 bp
* 46478: contig of 12884 bp in length
* 59362: gap of 100 bp
* 59462: contig of 15760 bp in length
* 75222: gap of 100 bp
* 75322: contig of 18050 bp in length
* 93372: gap of 100 bp
* 93472: contig of 23992 bp in length
* 117464: gap of 100 bp
* 117564: contig of 20152 bp in length
* 137716: gap of 100 bp
* 137816: contig of 26687 bp in length.
  Location/Qualifiers
    1. .164502
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="3"
      /map="3"
      /clone="RP11-170K4"
      /clone_lib="RP11-11 Human Male BAC"
  misc_feature
    1. .2917
      /note="assembly_fragment"
  misc_feature
    3018. .6310
      /note="assembly_fragment"
  misc_feature
    6411. .9447
      /note="assembly_fragment"
  misc_feature
    9548. .14470
      /note="assembly_fragment"
  misc_feature
    14571. .20528
      /note="assembly_fragment"
  misc_feature
    20629. .24667
      /note="assembly_fragment"

```

FEATURES

source

```

/note="assembly_fragment
clone_end:SP6
vector_side:right"
24768. .29167
/note="assembly_fragment"
29468. .36237
/note="assembly_fragment"
36338. .46377
/note="assembly_fragment"
46478. .59361
/note="assembly_fragment"
59462. .75221
/note="assembly_fragment"
75322. .93371
/note="assembly_fragment"
93472. .117463
/note="assembly_fragment"
117564. .137715
/note="assembly_fragment
clone_end:T7
vector_side:left"
137816. .164502
/note="assembly_fragment"

ORIGIN
Alignment Scores:
Pred. No.: 4.45e-163 Length: 164502
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-763-972-2 (1-360) x AC026960 (1-164502)
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgGlySerArgGlyValAlaLeu 20
Db 123728 ATGCTGTCCATTTGCTTCCTCCAGGGGAGAGAGCGGGAGCGGTGGAGCTGTG 123669
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
Db 123668 CTCCTGGAGGGAGGCTCCCGGGACATGGAGAAGGTGGACATGAATACATCAGGAACAA 123609
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerile 60
Db 123608 GGTCTCTGCCAGTTCCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 123549
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTyrGlyGlnThr 80
Db 123548 ATCTTTATCTTAGGGCTGCCACTAAATGGCAGCTGTCTTGTGGCACTCCTGGGGCCCAACC 123489
QY 81 LysArgTyrSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 123488 AAGCGCTGAGCTGGCCACCACCTATCTGGTGAACCTGTATGGTGGCCAGCTGCTTTTAT 123429
QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTyrProPheGlyGlu 120
Db 123428 GTGCTATTGCCCTTCCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 123369
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 123368 CTGCTCTCAAGCTGGTGCACCTCTCTGTCTTATATCAACCTTTACGGCAGCATCTGCTGTG 123309
QY 141 LeuTyrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 123308 CTGACCTGCATCTCTGTGCACCAAGTTCCTAGGTGTGTGCCACCACCTGTGTTCGTGCC 123249
QY 161 TyrArgThrArgArgHisAlaThrPleuGlyThrSerThrThrTyrAlaLeuValValLeu 180
Db 123248 TACCGAGCCGCGAGCATGCTGCTGGTGGCCACCAGCACCATCTGGGGCCCTGGTGTCTCTC 123189
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200

```

Db 123188 CAGCTGTCGCCACACTGGCTTCTCCACACGAGCTACATCATGATGATGTCG 123129
 Qy 201 TyrAspMetThrSerGlnGlnuAaPheAaArgLeuPheAlaTyrGlyIleValLeuThr 220
 Db 123128 TATGACATGACACAGCAGAGAAATTTGATCGGCTTTTGCCTACGCGATGTTCTGACA 123069
 Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 Db 123068 TTGCTGCTGCTTTCTTCCCTCCTTGGTCAATTTTGGTGTGCTATTCACTGATGTCAGGAG 123009
 Qy 241 ProAspGlnAlaArgGlyGluProHisGlnuAaPheArgGlnHisSerProSerGlnValHis 260
 Db 123008 CCTGATCAGCAGAGGAGAACCTTCATGAGACAGGACACAGCCGACCCAGGTCCAT 122949
 Qy 261 ProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrHis 280
 Db 122948 CCGGACCATCTACTGTGTGTGGCTCTTCCACCTCTGTTTGTGGCTTCCATATCAC 122899
 Qy 281 SerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 Db 122888 TCGCTCCTCTTACTCTACCATCTGCTTCTGCTTCTCAGACTGCGAGCTCTTGATGGC 122829
 Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
 Db 122828 AGCCAGTGTGGCCTTACAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCTCAACCC 122769
 Qy 321 SerProValLeuSerPheLeuGlyGlyValuAaArgValArgLeuGlnLeuLysLeuArg 340
 Db 122768 AGTCTCTGACTTCTTTTCAAGGGGGGCAAAATAGACTCAGGCTCTCCAGAACTGAGG 122709
 Qy 341 GlnAsnLysLeuGlyGluHisPheProAlaGlyValuAaArgCysProGlyLeuAsnArgSer 360
 Db 122708 CAGACCAAGTTGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGACAGATCT 122649
 RESULT 8
 AC092999 Homo sapiens 3 BAC RP11-170K4 (Roswell Park Cancer Institute Human
 LOCUS BAC Library) complete sequence.
 DEFINITION AC092999 Homo sapiens (human)
 ACCESSION AC092999.3 GI:19033407
 VERSION HTG.
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 167084)
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshman,P.R., Allen,C.,
 Alsbrooks,S.B., Amarantunga,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbaxia,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieval,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J.J., Foster,P., Frantz,P.,
 Gabis,A., Geo,J., Garcia,A., Garner,I., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,J., Li,J., Li,J., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
 Louisleged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,

Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
 Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
 Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N.,
 Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
 Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
 Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S.,
 Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E.,
 Sonait,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
 Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
 Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
 Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
 Warren,R., Washington,C., Watlington,S., Williams,G.,
 Williamson,A., Wleczkyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
 Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
 Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 167084)
 Worley,K.C.
 Direct Submission
 Submitted (09-AUG-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 167084)
 Worley,K.C.
 Direct Submission
 Submitted (01-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 167084)
 Worley,K.C.
 Direct Submission
 Submitted (29-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 5 (bases 1 to 167084)
 Worley,K.C.
 Direct Submission
 Submitted (30-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Mar 1, 2002 this sequence version replaced gi:18139250.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the
 Features listing.

ANNOTATION OF FEATURES:
 STSs are identified using ePCR (Genome Res. 7:541-550) searches
 of a local database that includes entries from dbSTS, GDB, and
 local mapping efforts.
 Repeats are identified using RepeatMasker (A. Smit and P. Green,
 unpublished.) for Human and Mouse sequences.
 Genes and Region of sequence similarity are identified by BLAST
 (Nuc. Acids Res 25:3389-3402) similarity (expect < 1e-34) to the
 EST and cDNA sequences. Genes demonstrate at least two exons
 flanked by consensus splice sites that maintained sequence
 continuity across the splice junctions. Sequences that are not
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
 standard of double strand coverage with a minimum of 2 clones and 2
 reads with no ambiguities or 2 chemistries with a minimum of 2
 clones and 3 reads with no ambiguities. If the sequence quality for
 a region does not meet this standard, it will be indicated in the
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality


```

Db      161581  TTGCTGGCTTCTTCTCCCTCTGCTGGTCAATTTGGTGTGCTATTCACTGATGTCAGGAG 161640
Qy      241    ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db      161641  CCTGATCAAGCCAGGAGGAGAACCTCATGAGCAGGCAACACAGCCGAGCCAGGTCAT 161700
Qy      261    ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
Db      161701  CCGACCACTCTACTCGGTGTGTGGCTCTTCACTCTCTGTTTGTGCTCCCTCCATATCAC 161760
Qy      281    SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
Db      161761  TCGCTCCTTCTACCTACCACTCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGC 161820
Qy      301    SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
Db      161821  AGCCAGTGTGGCCACAGATATGAGGCGCTCTGGTGTGTGAGCAGCTGCCCTCAACCC 161880
Qy      321    SerProValLeuSerPheLeuGlyGlyLeuGlyLeuAsnArgValArgLeuLeuGlnLysLeuArg 340
Db      161881  AGTCCTGTACTTCTTCAAGGGGGGCAAAATAGATCAGGCTCTCCACGAACTGAGG 161940
Qy      341    GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
Db      161941  CAGAACAGATTGGGTGAGCATCCAGCTGGGAGGAAGATGCCCAGGGTTGAACAGATCT 162000

RESULT 9
LOCUS   AC021773 Homo sapiens clone RP11-12N17, WORKING DRAFT SEQUENCE, 15 unordered
DEFINITION
ACCESSION AC021773
VERSION   AC021773.6 GI:10047750
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 163958)
Biren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome, clone RP11-12N17
Unpublished
2 (bases 1 to 163958)
Biren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavskiy,L., Boukhgaltier,B., Brown,A., Burkett,G., Castle,A.,
Choeple,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Rhoman,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.

Direct Submission
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 10, 2000 this sequence version replaced gi:7248967.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

```

----- Project Information
Center project name: L3302
Center clone name: 12_N17
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152766 bases at least Q40
Consensus quality: 158075 bases at least Q30
Consensus quality: 160495 bases at least Q20
Insert size: 174000; agarose-fp
Insert size: 162558; sum-of-ctnigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.9 in Q20 bases; sum-of-ctnigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 7254: contig of 7254 bp in length
* 7255 7354: gap of 100 bp
* 7355 8612: contig of 1258 bp in length
* 8613 8712: gap of 100 bp
* 8713 11463: contig of 2751 bp in length
* 11464 11563: gap of 100 bp
* 11564 14507: contig of 2944 bp in length
* 14508 14607: gap of 100 bp
* 14608 22452: contig of 7845 bp in length
* 22453 35086: contig of 12534 bp in length
* 35087 35186: gap of 100 bp
* 35187 44278: contig of 9092 bp in length
* 44279 44378: gap of 100 bp
* 44379 56786: contig of 12408 bp in length
* 56787 56886: gap of 100 bp
* 56887 65270: contig of 8384 bp in length
* 65271 65370: gap of 100 bp
* 65371 76394: contig of 11024 bp in length
* 76395 76494: gap of 100 bp
* 76495 100320: contig of 23826 bp in length
* 100321 100420: gap of 100 bp
* 100421 113127: contig of 12707 bp in length
* 113128 113227: gap of 100 bp
* 113228 126756: contig of 13529 bp in length
* 126757 126856: gap of 100 bp
* 126857 151061: contig of 24205 bp in length
* 151062 151161: gap of 100 bp
* 151162 163958: contig of 12797 bp in length.

Location/Qualifiers
1..163958
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-12N17"
/clone_lib="RPC1-11 Human Male BAC"
1..7254
/note="assembly_fragment
clone end:SP6
vector side:left"
7355..8612
/note="assembly_fragment"
8713..11463
/note="assembly_fragment"
11564..14507
/note="assembly_fragment"
14608..22452
/note="assembly_fragment"
22553..35086
/note="assembly_fragment"

misc_feature
1..7254
/note="assembly_fragment
clone end:SP6
vector side:left"
7355..8612
/note="assembly_fragment"
8713..11463
/note="assembly_fragment"
11564..14507
/note="assembly_fragment"
14608..22452
/note="assembly_fragment"
22553..35086
/note="assembly_fragment"

FEATURES
source
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

Db	98703	TATGACATGACGAGCAAGAGAAATTTTATCGGCTTTTTTGCTACGGCATAGTTC	TGTGACA	98762
Qy	221	LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyClnGlu	240	
Db	98763	TTGTCTGGCTTCTTTCCCTCCTTGCTCATTTTGGTGTGCTATTCACTGATGGTCAGGAG	98822	
Qy	241	ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis	260	
Db	98823	CCTGATCAAGCCAGAGGAGAACTCATGAGACAGCAACACAGCCGAGCGAGTCCAT	98882	
Qy	261	ProAspHisProThrGlyValTIpProLeuHisProLeuPheCysAlaLeuProTyrHis	280	
Db	98883	CGGACCATCCTACTGGTGTGTGGGCTCTTACCCCTCTGT-TTGTGCCCTTCCATATCAC	98942	
Qy	281	SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly	300	
Db	98943	TGCTCTCTTCTACTACCTCACCATCTGCTTCTTCTGCTTTCTCAGGACTGCCAGCT	99002	
Qy	301	Ser-GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPr	320	
Db	99003	AGCCCAAGTGTGGCTTACAGATATGAGAGCCCTCTGGTAGTGTGAGCAGCTGCC	99062	
Qy	320	oSerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuAr	340	
Db	99063	CAGTCTGTACTTCTTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCCAGAACTCAG	99122	
Qy	340	GGlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSe	360	
Db	99123	GCAGAAACAAGTTGGGTGAGCATCCAGCTGGGAGGAAGATGCCCAGGGTTGACAGATC	99182	
Qy	360	r	360	
Db	99183	T	99183	
RESULT 10				
AX921825				
LOCUS				
DEFINITION				
Sequence 165 from Patent WO02068649.				
ACCESSION				
AX921825				
VERSION				
AX921825.1				
KEYWORDS				
SOURCE				
Homo sapiens (human)				
ORGANISM				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE				
1				
AUTHORS				
JOURNAL				
FEATURES				
source				
Patent: WO 02068649-A 165 06-SEP-2002;				
Curagen Corporation (US)				
Location/Qualifiers				
1..1108				
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
ORIGIN				
Alignment Scores:				
Pred. No.:				
2,51e-162				
Length:				
1108				
Score:				
1896.00				
Matches:				
359				
Percent Similarity:				
99.45%				
Conservative:				
0				
Best Local Similarity:				
99.45%				
Mismatches:				
1				
Query Match:				
97.93%				
Indels:				
2				
DB:				
Gaps:				
0				

Db 72 CTCCTGGAGGAGCTCCGGGACATGGAGAGGTGGACATGATATACATCACAGGAACA 131
 QY 41 GlyLeuCysGlnPheSerGluValLeuValTyrLeuSerLeuAlaTyrSerile 60
 Db 132 GGTCTCTGCCAGTCTCAGAGAGTACAGGAGGTACCTCTCCCTGGGCTACAGTATC 191
 QY 61 IlePheLeuLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
 Db 192 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGCCAAACC 251
 QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
 Db 252 AAGGCTGGAGCTGGCCACCACTATCTGTGTGAACCTGATGGTGGCGACCTGCTTTAT 311
 QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120
 Db 312 GTGCTATTGCCCTTCTCATCATCACTACTCATAGATGACAGTGGCCCTTCGGGGAG 371
 QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerileLeuLeu 140
 Db 372 CTGCTCTGCAAGCTGGTGCACTCTCTGTCTTATATCAACCTTTACGGCAGCATCTGCTG 431
 QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 Db 432 CTGACCTGCACTCTGTGGCACCAGTCTCTAGGTGTGTGCCACCACCTGTGTGCTGCC 491
 QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeu 180
 Db 492 TACGGGACCCGAGGCATGCTGGCTGGCACACAGCACCACTGGGCGCTGTGTGCTCTC 551
 QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
 Db 552 CAGCTGTGCTGCCACACTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 611
 QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
 Db 612 TATGACATGACCCAGCAAGAAATTTGATCGGCTTTTGGCTACGCGCATGTCTGACA 671
 QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 Db 672 TTGTCTGGCTTCTTCTCCCTCTGGTCACTTTGGTGTGTA-TTCACTGATGTGTGAGG 730
 QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
 Db 731 CCGTATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACAGCCCGAGCCAGGTCCAT 790
 QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
 Db 791 CCGGACCATCTACTGTGTGTGGCTCTTCCACCTCTGTGTGTGGCTTCCATATCAC 850
 QY 281 SerLeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 Db 851 TCGCTCTCTTACTCACTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 910
 QY 301 Ser-GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnLeuProGlnPr 320
 Db 911 AGCCAGGTGGCTTACAGATATGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 970
 QY 320 oSerProValLeuSerPheHisGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
 Db 971 CAGTCTCTGTATCTTCTTCAAGGGGGGGGCAAAATAGAGTCAAGGTCTCTCCAGAACTGAG 1030
 QY 340 gGlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgse 360
 Db 1031 GCAGAACAAAGTTGGTGAGCATCCAGCTGGGAGGAGAGATGCCCGGGTTGAACAGATC 1090
 QY 360 r 360
 Db 1091 t 1091
 RESULT 11
 AF411108
 LOCUS AF411108 1019 bp DNA linear PRI 01-NOV-2001

DEFINITION Homo sapiens GPR79 pseudogene, complete sequence.
 ACCESSION AF411108
 VERSION AF411108.1 GI:16566321
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 1019)
 AUTHORS Lee, D.K., Nguyen, T., Lynch, K.R., Cheng, R., Vanti, W.B., Arkhitko, O., Lewis, T., Evans, J.F., George, S.R. and O'Dowd, B.F.
 TITLE Discovery and mapping of ten novel G protein-coupled receptor genes
 JOURNAL Gene 275 (1), 83-91 (2001)
 MEDLINE 21458557
 PUBMED 11574155
 REFERENCE 2 (bases 1 to 1019)
 AUTHORS Lee, D.K., Nguyen, T., Lynch, K.R., Cheng, R., Vanti, W.B., Arkhitko, O., Lewis, T., Evans, J.F., George, S.R. and O'Dowd, B.F.
 TITLE Direct Submission
 JOURNAL Submitted (17-AUG-2001) Department of Pharmacology, University of Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A5, Canada
 FEATURES
 source
 1..1019
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene
 1..1019
 /gene="GPR79"
 /pseudo
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,278-154 Length: 1019
 Score: 1806.00 Matches: 339
 Percent Similarity: 99.71% Conservative: 0
 Best Local Similarity: 99.71% Mismatches: 0
 Query Match: 93.29% Indels: 1
 DB: Gaps: 0
 US-10-763-972-2 (1-360) x AF411108 (1-1019)
 QY 1 MetLeuSerileLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
 Db 1 ATGCTGTCCATTTTCTTCCAGGGAAGCAGAGCGGAGCGCTGCTGGAGCTCTG 60
 QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
 Db 61 CTCTGGAGGAGGCTCCCGGACATGAGAGGTGGACATGATATACATCACAGGAACA 120
 QY 41 GlyLeuCysGlnPheSerGlyLysValTyrLeuSerLeuAlaTyrSerile 60
 Db 121 GGTCTCTGCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCTCTGCTCAGTATC 180
 QY 61 IlePheLeuLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
 Db 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGCCAAACC 240
 QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
 Db 241 AAGCCCTGGAGCTGTGCCACCACCTATCTGTGTGAACCTGATGGTGGCGACCTGCTTAT 300
 QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120
 Db 301 GTGCTATTGCCCTTCTCTCATCATCACTACTCATAGTACAGGTGGCCCTTCGGGGAG 360
 QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerileLeuLeu 140
 Db 361 CTGCTCTGCAAGCTGGTGCACTTCTCTGTCTTATATCAACCTTTACGGCAGCATCTGCTG 420
 QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 Db 421 CTGACCTGCACTCTCTGTGACAGTTCCTAGGTGTGTGCCACCCACTGTGTGTGCTGCC 480


```

Db      721 CCTTTTACCCCTCTGTTTGTGGCTTCCATATACCTCGTCTCTTACCTCACCACCTG 780
QY      289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
Db      781 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCTTACAGATATG 840
QY      309 GluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPheLeysGly 328
Db      841 GAGGCTCTGTGTGAGTGTGAGCAGCTGCTCAACCCAGTCTCTGCTGCTCTTCTTCAAGGG 900
QY      329 GlyLeuAsnArgValArgLeuGlnLeuGlnLeuArgGlnAsnLeuGlyGluHisPro 348
Db      901 GCGAATAATAGAGTCAGCTCTCTCGAAGACTGAGGAGAGACAGTTGGGTGAGCATCCA 960
QY      349 AlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
Db      961 GCTGGAGAGAGATGCCAGGCTTGAACAGATCT 996

RESULT 13
LOCUS   BD187417
DEFINITION Novel polypeptide.
ACCESSION BD187417
VERSION   BD187417.1 GI:32997156
KEYWORDS JP 2003009885-A/2.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE    Fido, M.D.
JOURNAL  1 (bases 1 to 1020)
COMMENT  Novel polypeptide
        Patent: JP 2003009885-A 2 14-JAN-2003;
        Pfizer Ltd (EP(GB) only), Pfizer Inc (US JP EP except GB)
        OS Homo sapiens
        PN JP 2003009885-A/2
        PD 14-JAN-2003
        PF 17-DEC-2001 JP 2001382707
        PR 18-DEC-2000 GB 0030855.1, 17-JAN-2001 GB 0101222.8 PI
        mark david fido
        CC

FEATURES
source   Location/Qualifiers
          1..1020
          /organism="Homo sapiens"
          /mol_type="genomic DNA"
          /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2.63e-152 Length: 1020
Score: 1785.00 Matches: 337
Percent Similarity: 99.12% Conservative: 1
Best Local Similarity: 98.83% Mismatches: 1
Query Match: 92.20% Indels: 2
DB: 6 Gaps: 0

US-10-763-972-2 (1-360) x BD187417 (1-1020)

QY      1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
Db      1 ATGCTGTCCATTTGCTTCTCTCCAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY      21 LeuLeuGluGlyAlaSerArgMetGluCysValAspMetAsnThrSerGlnGluGln 40
Db      61 CTCCTGGAGGGAGCTCCCGGACATGAGAGAGTGGACATGAATACATCACAGGACAA 120
QY      41 GlyLeuCysGlnPheSerGluGlySerGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db      121 GGTCTCTCCAGTCTCTCAGAGAGTACAGCAAGTCTACTCTCTCCCTGGGCTACAGTATC 180
QY      61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThr 80

```

```

Db      181 ATCTTTATCCCTAGGCTGCCACTAAATAGCACTGCTCTTGTGGCACTCTCTGGGCGCAACC 240
QY      81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db      241 AAGCGCTGGAGCTGTGGCCACCACCTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTAT 300
QY      101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120
Db      301 GTGCTATTGGCCCTTCTCATCATCACCTACTACTAGATGACAGGTGGCCCTTTCGGGAG 360
QY      121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db      361 CTGCTCTGCAAGCTGGTGCACCTTCTCTTATATCAACCTTTACGGGAGCATCTCTGCTG 420
QY      141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db      421 CTGACCTGCATCTCTGTGCACCCAGTTCCTAGGTGTGGCACCCCACTGTGTGCTGCC 480
QY      161 TyrArgThrArgGlnHisAlaTyrLeuGlyThrSerThrThrTrpAlaLeuValLeu 180
Db      481 TACCGAGCCCGAGGAGCATGCTGTGGTGGCACCACTGGGCCCTCTGGTGGTCTC 540
QY      181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
Db      541 CAGCTGTGCTGCCACACTGGCCCTTCTCCACACGGAATACTCAATGGCCAGATCTGG 600
QY      201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db      601 TATGACATGACAGGAGCAAGAAATTTTATCGGCTTTTGTCTCGGCAATAGTCTTGACA 660
QY      221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGln 240
Db      661 TTGTCGTGGCTTTCTTCCCTCCCTCTGCTATTTGTGTGCTATTCATCTGCTCAGG 720
QY      240 LuProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH 260
Db      721 AGCCTGTATCAAGCCAGAGAGAACTCATGAGGAGGAGCAACACAGCCGAGGAGTCC 780
QY      260 IspAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrH 280
Db      781 ATCCCGACCATCTACTGCTGTGGCTCTTCCCTCTGCTTTGTGCTCTCCATATC 840
QY      280 IspSerLeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
Db      841 ACTCGCTCTCTTACTCTCACCATCTGCTTTCTGCTTCTCAGGACTGCCAGCTCTTGATG 900
QY      300 LysSerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320
Db      901 GCACCCAGTGTGGCTTACAGATATGGAGGCTCTGGTGTGAGTGTGAGGAGTGGCTCAAC 960
QY      320 roSerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeu 339
Db      961 CCAGTCTCTGCTACTTTCTTCAAGGGGCGCAAAATAGAGTCAAGTCTCTCCAGAACTG 1019

RESULT 14
LOCUS   AX458308
DEFINITION Sequence 3 from Patent EP1215214.
ACCESSION AX458308
VERSION   AX458308.1 GI:21725002
KEYWORDS Homo sapiens (human)
SOURCE   Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE    Fido, M.D.
JOURNAL  1
        Novel polypeptide
        Patent: EP 1215214-A 3 19-JUN-2002;
        Pfizer Limited (GB); PFIZER INC. (US)
        Location/Qualifiers
        1..1020

```

```
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,63e-152 Length: 1020
Score: 1785.00 Matches: 337
Percent Similarity: 99.12% Conservativeness: 1
Best Local Similarity: 98.83% Mismatches: 1
Query Match: 92.20% Indels: 2
DB: 6 Gaps: 0

US-10-763-972-2 (1-360) x AX458308 (1-1020)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
Db 1 ATGCTGTCCATTTTTCCTTCCTCAGGGGAAGCAGAGCGGAGCGTGTGGAGCTCTG 60
Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
Db 61 CTCTGTGGAGGAGCTCCCGGACATGGAGAGGTGGACATGAATACATCACAGGAACA 120
Qy 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerile 60
Db 121 GGTCTCTGCCAGTCTCTCAGAGAAGTACACAGAAAGTCTACCTCTCCCTGGCTACAGTATC 180
Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTyrGlyGlnThr 80
Db 181 ATCTTTATCTTAGGGCTGCCACATAAATGGCACATGTCTTGTGGCACCTCTGGGGCCAAACC 240
Qy 81 LysArgTyrSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 241 AAGGCTGTGAGCTGTGCCACCACTATCTGTGAACCTGATGTGGCCGACCTGCTTTAT 300
Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTyrProPheGlyGlu 120
Db 301 GTGCTATTGGCCCTTCTCTCATCATCACTACTCATAGTACAGAGTGGCCCTTCGGGGAG 360
Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerileLeuLeu 140
Db 361 CTGCTCTGCAAGCTGTGGTCACTCTCTCTTATATCAACCTTTACGGCAGCATCTGCTG 420
Qy 141 LeuThrCysLeuSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 421 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGGCACCACTGTGTCTGCTGCC 480
Qy 161 TyrArgThrArgHisAlaTyrLeuGlyThrSerThrThrThrAlaLeuValLeuLeu 180
Db 481 TACGGAGCCCGAGCGATGCTGGCTGGCAGCACCACTGGGGCCCTGTGGTGGCTC 540
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
Db 541 CAGCTGTGCCACACATGGCTTCTCCACACGAGCTTACATCAATGGCCAGATGATCTGG 600
Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 601 TATGACATGACACGCAAGAGAAATTTTGATCGGCTTTTGGCTACGGCATAGTTCTTGACA 660
Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG 240
Db 661 TTGTCTGGCTTCTTTTCCCTCTCTGGTCAATTTTGTGTGCTATTCACCTGATGGTTCAGG 720
Qy 240 LuProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH 260
Db 721 AGCCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGGTCC 780
Qy 260 IsProAsnHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrH 280
Db 781 ATCCGACCACTCTACTGGTGTGGCTCTTCACTCCCTCTGTTTGTGGCTTCCATATC 840
Qy 280 IsSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
```

```
Db      417 CTGCTCTCAAGCTGGTGCACTTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 476
Qy      141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
      |||
Db      477 CTGACCTGCATCTCTGTGCACAGTCTCTAGTGTGTGGCACCACCTGTGTGCTGCC 536
Qy      161 TyrArgThrAspArgHisAlaTyrLeuGlyThrSerThrThrTyrAlaLeuValValLeu 180
      |||
Db      537 TACCGGACCCGAGGATGCTGCTGGGCACGACCCACCTGGGGCCCTGGTGGTCTC 596
Qy      181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
      |||
Db      597 CAGCTGTGCCCCACACTGGCCCTTCTCCACACGGACTACATCAATGGCCAGATGATCTGG 656
Qy      201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
      |||
Db      657 TATGACATGACGAGCAAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCGACA 716
Qy      221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG 240
      |||
Db      717 TTGTCTGGCTTTCTTTCCCCCTCTTGGTCACTTTTGGTGTCTATTCACTGATGGTCAGG 776
Qy      240 LuProAspGlnAlaArgGlyGlnProHisGluAspArgGlnHisSerProSerGlnValH 260
      |||
Db      777 AGCCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCCCGAGCCAGGTCC 836
Qy      260 isProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrH 280
      |||
Db      837 ATCCGGACCATCTACTGGTGTGGCTCTTCACTCTGTGTTTGGCCCTTCCATATC 896
Qy      280 isSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
      |||
Db      897 ACTGCTCTCTTCTACCTCACCATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGATG 956
Qy      300 LySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320
      |||
Db      957 GCACCCAGTGTGGCTACAGATATGAGGGCTCTGGTGAGTGTGAGCAGCTGCTCCAC 1016
Qy      320 roSerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeu 339
      |||
Db      1017 CCAGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCCAGAAACTG 1075
```

Search completed: November 10, 2004, 14:48:30
Job time : 5424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 10, 2004, 11:49:21 ; Search time 111 Seconds
(without alignments)

2305.261 Million cell updates/sec

Title: US-10-763-972-2

Perfect score: 1936

Sequence: 1 MSLILPSRSGSRGAL.....QNKLGHPAGKRCFCPLNRS 360

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			
Xgapext 0.5			
Fgapop 6.0			
Fgapext 7.0			
Delop 6.0			
Delext 7.0			

Searched:

824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10763972/runat 04112004 093314 8810/app query.fasta_1.519
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOCPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10763972 /cgn2_1_69 @runat 04112004 093314 8810 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DGPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Match	Length	ID	Description
1	445.5	23.0	1429	4	US-09-016-434-1068
2	445.5	23.0	1429	4	US-09-077-173D-1
3	442	22.8	2025	4	US-09-016-434-1482
4	442	22.8	2025	4	US-09-814-915A-74
5	420.5	21.7	984	3	US-08-513-974B-41
6	420.5	21.7	984	4	US-09-461-436B-41
7	420.5	21.7	1020	3	US-08-513-974B-370
8	420	21.7	1842	1	US-08-442-134A-1
9	420	21.7	1842	1	US-08-444-581B-1
10	420	21.7	1842	1	US-08-446-088A-1
11	413	21.3	984	3	US-08-513-974B-57
12	413	21.3	984	4	US-09-461-436B-57

13	21.3	413	1023	3	US-08-513-974B-379	Sequence 379, Appl
14	411	21.2	1571	4	US-09-016-434-1108	Sequence 1108, Ap
15	402	20.8	984	3	US-08-459-046-1	Sequence 1, Appli
16	402	20.8	984	4	US-09-102-710B-1	Sequence 1, Appli
17	313.5	16.2	3055	4	US-09-016-434-1456	Sequence 1456, Ap
18	275	14.2	1140	4	US-09-016-434-750	Sequence 750, Appl
19	275	14.2	1301	2	US-08-467-948A-7	Sequence 7, Appli
20	275	14.2	1301	3	US-08-467-947A-7	Sequence 7, Appli
21	268.5	13.9	1346	4	US-09-761-962A-12	Sequence 12, Appl
22	268	13.8	2160	3	US-08-188-275A-1	Sequence 1, Appli
23	268	13.8	2162	3	US-09-351-198-1	Sequence 1, Appli
24	268	13.8	2162	3	US-09-113-436-1	Sequence 1, Appli
25	268	13.8	2162	4	US-09-016-434-1379	Sequence 1379, Ap
26	268	13.8	2162	4	US-09-355-709C-7	Sequence 7, Appli
27	267.5	13.8	1981	3	US-08-387-707-15	Sequence 15, Appl
28	267.5	13.8	1981	4	US-08-405-271A-15	Sequence 15, Appl
29	267	13.8	1610	3	US-08-889-108-7	Sequence 7, Appli
30	267	13.8	1610	5	PCT-US94-10358-7	Sequence 7, Appli
31	266.5	13.8	1900	4	US-09-016-434-1484	Sequence 1484, Ap
32	266.5	13.8	1901	1	US-08-153-848-43	Sequence 43, Appl
33	266.5	13.8	1901	3	US-09-299-843A-43	Sequence 43, Appl
34	266.5	13.8	1901	3	US-09-088-337B-43	Sequence 43, Appl
35	266.5	13.8	1901	5	PCT-US93-11153-43	Sequence 43, Appl
36	266.5	13.8	2453	5	PCT-US95-07180-1	Sequence 1, Appli
37	266	13.7	1020	4	US-09-170-496D-181	Sequence 181, App
38	265	13.7	1020	4	US-09-170-496D-31	Sequence 31, Appl
39	264.5	13.7	1729	4	US-09-761-962A-9	Sequence 9, Appli
40	262.5	13.6	1205	1	US-08-417-103-13	Sequence 13, Appl
41	262.5	13.6	1334	4	US-09-761-962A-3	Sequence 3, Appli
42	262.5	13.6	1365	4	US-09-761-962A-11	Sequence 11, Appl
43	262.5	13.6	1423	4	US-09-761-962A-1	Sequence 1, Appli
44	262.5	13.6	1542	4	US-09-761-962A-4	Sequence 4, Appli
45	262.5	13.6	1610	4	US-09-761-962A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-016-434-1068
; Sequence 1068, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; FILING APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555

```

; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1068:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1429 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1124904
; US-09-016-434-1068

Alignment Scores:
Pred. No.: 1,24e-37 Length: 1429
Score: 445.50 Matches: 123
Percent Similarity: 50.00% Conservative: 47
Best Local Similarity: 36.18% Mismatches: 131
Query Match: 23.01% Indels: 40
DB: Gaps: 9

US-10-763-972-2 (1-360) x US-09-016-434-1068 (1-1429)
QY 6 LeuPro---SerArgGlySerArgSerGlySerArgGlyAlaLeuLeuGly 24
Db 160 CTTCCCATCTCCAGGGGGCCATGCCAGTACAGAGTCCCTCTGTGAGATCCCTAGGC 219
QY 25 AlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGlnGlyLeuCysGln 44
Db 220 CTCAGC-----CCAGGTCCTGGCAGCAGTGGAGTGGAGCTGTTGG 264
QY 45 PheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerLeilellePheLeu 64
Db 265 TTTGATGAGGATTCAAGTTCACTGCTGCTGTGAGTATGAGTGTCTTTGTGCTG 324
QY 65 GlyLeuProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThrLysArgTrpSer 84
Db 325 GGCCTGGGCTTAAAGCCCAACCTATGGCTCTTCATCTTCGCGCTCCGACCTGGAT 384
QY 85 CysAlaThrTyrLysGlnValAsnLeuMetValAlaAspLeuLeuTyrValLeu---Leu 103
Db 385 GCACCGCCACCTACTATTCACCTGCATGTGCACACCTGTATGTGTGCTGCTG 444
QY 104 ProPheLeuIlelleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCys 123
Db 445 CCCACCTCATCTACTATATGACGCCACACCACTGGCCCTTTGGCACTGAGATGCG 504
QY 124 LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIlellePheLeuCys 143
Db 505 AAGTTCGTCGCTTCTTTCTATGAGACCTCTACTGCGAGTGTCCCTTTCCCTACCTGC 564
QY 144 IleSerValHisGlnLeuGlyValCysHisProLeuCysSerLeuProTyrArgThr 163
Db 565 ATCAGCGTGACCGCTACTCGGCATCTGCCACCACTTCGGGCACCTACGCTGGGCGCG 624
QY 164 ArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeu 183
Db 625 CCTCGCCTCGAGCGCTTCTCTGCTGCGCAGTGTGGTGTGCTGAGCGCTCGCTCGTG 684
QY 184 ProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMet 203
Db 685 CCCAACCTGTTCTTGTGCACACCAACAGCAAAAGGACCCAGCTGCTGTCATGACACC 744
QY 204 ThrSerGlnGluAsnPhaAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGly 223
Db 745 ACTCGGCTGAGAGTGTGACCA-CTATGTGCA-----CTTCAGCTCGGCGGT 791
QY 224 -----PheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGly 238
Db 792 CATGGGCTGCTCTTTGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
QY 239 GlnGluProAspGlnAlaArgGly-----GluProHis.sGluAsp 251
Db 852 TCGTCGCTGTATCAGCCCTTGGCCAGGCTCTGCACAGTGTCTTCTTCGCTGCTGCTG 902
b

```


STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/513,974B
 FILING DATE: 14-SEP-1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP95/01599
 FILING DATE: 10-AUG-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-093989
 FILING DATE: 19-AUG-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-057186
 FILING DATE: 16-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-007177
 FILING DATE: 20-JAN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-326611
 FILING DATE: 28-DEC-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-270017
 FILING DATE: 02-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-236357
 FILING DATE: 30-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-236356
 FILING DATE: 30-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-189274
 FILING DATE: 11-AUG-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-189273
 FILING DATE: 11-AUG-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-189272
 FILING DATE: 11-AUG-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Resnick, David S.
 REGISTRATION NUMBER: 34,235
 REFERENCE/DOCKET NUMBER: 45753
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-523-3400
 TELEFAX: 617-523-6440
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 984 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-513-974B-41

Alignment Scores:
 Pred. No.: 3,18e-35 Length: 984
 Score: 420.50 Matches: 117
 Percent Similarity: 45.05% Conservative: 47
 Best Local Similarity: 32.14% Mismatches: 125
 Query March: 21.72% Indels: 76
 DB: 3 Gaps: 9

US-10-763-972-2 (1-360) x US-08-513-974B-41 (1-984)

QY 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeu-----CysGln 44
 |||||:|||||

Db	1	ATGGAGCAG---GACAATGGCACCATCCAGGCTCCAGGCTTGCGGCCACACCCACCTGGCGTC	57
QY	45	PheserGluLysTyrLysGlnValTyrLysSerLeuAlaTyrSerIleIlePheIleLeu	64
Db	58	TACCGTGAGGATTTCAAGCGACTCTGCTAAACCCCGGTATCTCGGTGGTCTGGTGTCT	117
QY	65	GlyLeuProLeuAsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArgTrpSer	84
Db	118	GGCTGGCCACTGAACATCTGGTCAATCCAGCATCTGGCATCCCGCGGACCGCTGACC	177
QY	85	CysAlaThrThrLeuValAsnMetValAlaLeuLeuTyrValLeu---Leu	103
Db	178	CGTCCGCTGTGTACACCCCTGAACCTGGCATCGGCGGACCTGATGATCGCTGTCTACTA	237
QY	104	ProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCys	123
Db	238	CCCTTACTTATCTATACTACGAGGGGACCACTGGCCCTTCGGAGACCTGGCTGC	297
QY	124	LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCys	143
Db	298	CGCTTTGTACGCTTCTCTATGCCAATCTACATGGCAGCATCTGCTTCTCCTACCTGC	357
QY	144	IleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---	162
Db	358	ATTAGCTTCCAGCGCTACCTGGGCATCTGCCACCCCTGGCTTCTGGCACAAGGTGGA	417
QY	163	ThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeu	182
Db	418	GGTCCCGTGTGCTTGGTAGTGTGGAGTCGTGGTGGCTGGCTGCACAGCCAGTGC	477
QY	183	LeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAsp	202
Db	478	CTGCCACGGCAGTCTTTGCTGCCAGCGCATCCAGCGCAACCCGACCTGTGTGCTACGAC	537
QY	203	MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSer	222
Db	538	CTGAGCCCAACCATCTCTGCTACTGCTACCTGCCCTATGATGATGCTGCTCAGGTGCTC	597
QY	223	GlyPheLeu--SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProA	242
Db	598	GGCTTCTTGTGCTGCTTCTATAGCTTACTGGCTTGTATTGTCG-----	641
QY	242	spGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHisProA	262
Db	642	-----CATGGCCCGCGCTGTGTGCGCAG	666
QY	262	spHisProThr-----	265
Db	667	GATGCCCCAGCAGGCTCTGTGCCCCAAGAGCGGCGCAGCGGCTCGTATGGCTGTG	726
QY	266	-----GlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeu	284
Db	727	GTGGTGGCAGCTGCTCTTTGCCATCAGCTT---CCTGCCCTTCCACATCACCAGACAGCC	783
QY	284	euProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysG	304
Db	784	TACTTGGCTGTGGCTTCCAGCGCGGTGCTCTTGGCCCTGTGCTGGAGACCTTCGCTGCT	843
QY	304	lyLeuGlnAspMetGluAlaSerGlyGluCysGluGln-----	316
Db	844	GCCTACAAGGACCTCGGCCCTTGGCCAGTGTCAACAGTGTTCGTGACCCCCATTCTTTC	903
QY	317	-----LeuProGlnProSerProValLeuSerPheLysGlyGlyLysA	331
Db	904	TACTTCACACACAGAAGTTCCGGCGGC-AAACCCACAGATCTCTTACAGAGCTCACAG-	961
QY	331	snArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProAlaGlyA	351
Db	962	-----CCAAGTGGCA	971
QY	351	rgLysArg	353
Db	972	GAGGCAGA	979

RESULT 6

US-09-461-436B-41
 ; Sequence 41, Application US/09461436B
 ; Patent No. 6538107
 ; GENERAL INFORMATION:
 ; APPLICANT: Shuji Hinuma
 ; Yasuaki Ito
 ; Eyo Fujii

TITLE OF INVENTION: G Protein Coupled Receptor Protein,
 ; Production, And Use Thereof

NUMBER OF SEQUENCES: 61

CORRESPONDENCE ADDRESS: 61

ADDRESSEE: Edwards & Angell, LLP

STREET: 101 Federal Street

CITY: BOSTON

STATE: MA

COUNTRY: USA

ZIP: 02209

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/461,436B

FILING DATE: 14-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/513,974

FILING DATE: 14-SEP-1995

APPLICATION NUMBER: PCT/JP95/01599

FILING DATE: 10-AUG-1995

APPLICATION NUMBER: 7-093989

FILING DATE: 19-APR-1995

APPLICATION NUMBER: 7-057186

FILING DATE: 16-MAR-1995

APPLICATION NUMBER: 7-007177

FILING DATE: 20-JAN-1995

APPLICATION NUMBER: 6-326611

FILING DATE: 28-DEC-1994

APPLICATION NUMBER: 6-270017

FILING DATE: 02-NOV-1994

APPLICATION NUMBER: 6-236357

FILING DATE: 30-SEP-1994

APPLICATION NUMBER: 6-236356

FILING DATE: 30-SEP-1994

APPLICATION NUMBER: 6-189274

FILING DATE: 11-AUG-1994

APPLICATION NUMBER: 6-189273

FILING DATE: 11-AUG-1994

APPLICATION NUMBER: 6-189272

FILING DATE: 11-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: CONLIN, DAVID G.

REGISTRATION NUMBER: <Unknown>

REFERENCE/DOCKET NUMBER: 45753 DIV2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-439-4444

TELEFAX: 617-439-4170

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 984 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 41:

US-09-461-436B-41

Alignment Scores:

Pred. No.: 3.18e-35 Length: 984

Score: 420.50 Matches: 117

Percent Similarity: 45.05% Conservative: 47
 Best Local Similarity: 32.14% Mismatches: 125
 Query Match: 21.72% Indels: 76
 DB: 4 Gaps: 9
 US-10-763-972-2 (1-360) x US-09-461-436B-41 (1-984)
 QY 29 MetGluValAspMetAsnThrSerGlnGluGlnGlyLeu-----CysGln 44
 Db 1 ATGGAGCAG--GACAATGGCACCATCCAGGTCAGGCTGCCGCCACACCTGCCTC 57
 QY 45 PheSerGluValTyLysGlnValTyLeuSerLeuAlaTyLysSerIlePheLeu 64
 Db 58 TACCGTGAGGATTCAAGCGACTGCTGCTAACCCCGGTATATCTCGTGGTGGTGGTGC 117
 QY 65 GlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSer 84
 Db 118 GGCCTGCCACTGAACATCTGCCTCATTCGCCAGATCTGCCGATCCGCCGCCACCTGACC 177
 QY 85 CysAlaThrThrTyLeuValAsnLeuMetValAlaAspLeuLeuTyValLeu---Leu 103
 Db 178 CGTTCGGCTGTGTACACCTGACCTGGCACTGGCGGACCTGATGATGCTGTCTACTA 237
 QY 104 PropheLeuIleThrTySerLeuAspArgTrpPropheGlyGluLeuLeuCys 123
 Db 238 CCCCTACTTATCTATAACTAGCCAGAGGGGACCACTGGCCCTTCGGAGACCTGCCTGC 297
 QY 124 LysLeuValHisPheLeuPheTyLysLeuTyLysSerIleAsnLeuTyGlySerIleLeuLeuThrCys 143
 Db 298 CGCTTTGTACGCTTCTCTTATGCCAATCTACATGACGACGATCTCTTCTTCTACCTGC 357
 QY 144 IleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyArg--- 162
 Db 358 ATTAGCTCCAGCGTACTGGGATCTGCCACCCCTGGCTCTCTGCCACAAGCGTGA 417
 QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeu 182
 Db 418 GGTCCCGCTGCTGTGGT 477
 QY 183 LeuProThrLeuAlaPheSerHisThrAspTyLysAsnGlyGlnMetIleTrpTyArg 202
 Db 478 CTGCCACGGCAGTCTTTGCTGCCACAGCATCCAGCGCAACCCGACCTGTGTGTACGAC 537
 QY 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyLysGlyIleValLeuThrLeuSer 222
 Db 538 CTGAGCCCAACCCATCTCTACTGCTACCTGCCCTATGTATGTATGTATGTATGTATGTAT 597
 QY 223 GlyPheLeu--SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProA 242
 Db 598 GGCCTTCTGTGCTGCTTCTATAGCCTTACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 641
 QY 242 spGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHisProA 262
 Db 642 -----CATGGCCCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 666
 QY 262 spHisProThr----- 265
 Db 667 GATGCCCGCAGCAGGTCTGTGCCCAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 726
 QY 266 -----GlyValTrpProLeuHisProLeuPheCysAlaLeuProTyTrpHisSerLeuLeu 284
 Db 727 GTGGTGGCAGTGTCTTTTGGCCTCAGCTT---CCTGCCCTTCCACATCACCACAGACGCC 783
 QY 284 euProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysG 304
 Db 784 TACTTGGCTGTGCGCTCCACGCCCGGTGTCTTGTGCCCTGTGTGTGTGTGTGTGTGTGTGT 843
 QY 304 LyLeuGlnAspMetGluAlaSerGlyGluCysGluGln----- 316
 Db 844 GCCTACAAAGGCACCTCGGCCCTTCGCCAGTGTCAACAGTGTCTGTGGAGCCCTTCTTCTTC 903
 QY 317 -----LeuProGlnProSerProValLeuSerPheLysGlyGlyLysA 331
 Db 317 ----- 331

Db 904 TACTTCACACAGAGTTCCGGCGC-AACCCACGATCTCTACAGAGGCTCACAG- 961
Qy 331 snArgValArgLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProAlaGlyA 351
Db 962 -----CCAAGTGGA 971
Qy 351 xGlysaG 353
Db 972 GAGCAGA 979

RESULT 7
US-08-513-974B-370
Sequence 370, Application US/08513974B
Patent No. 6114139
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ohtaki, Tetsuya
APPLICANT: Fukusumi, Shoji
APPLICANT: Ohgi, Kazuhiro
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
NUMBER OF SEQUENCES: 380
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513.974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 26-DEC-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994

ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 370:
SEQUENCE CHARACTERISTICS:
LENGTH: 1020 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 28..1011
US-08-513-974B-370
Alignment Scores:
Pred. No.: 3,37e-35 Length: 1020
Score: 420.50 Matches: 117
Percent Similarity: 45.05% Conservative: 47
Best Local Similarity: 32.14% Mismatches: 125
Query Match: 21.72% Indels: 76
DB: 3 Gaps: 9
US-10-763-972-2 (1-360) x US-08-513-974B-370 (1-1020)
Qy 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeu-----CysGln 44
Db 28 ATGGAGCAG--GACAATGGCACCATCCAGGCTCCAGGCTGCCGCCACACACCTGGCTC 84
Qy 45 PheSerGluLysTyrlsGlnValTyrlsLeuSerLeuAlaTyrlsSerllePheIleLeu 64
Db 85 TACCCTGAGGATTCAGAGGACTGCTGCTAACCCCGGTATCTCGTGGTGGTGGTGC 144
Qy 65 GlyLeuProLeuAsnGlyThrValleuTrpHisSerTrpGlyGlnThrLysArgTrpSer 84
Db 145 GGCTTGCCACTGACATCTCGTCTATTGCCAGATCTCCGCACTCCCGCGGACCTGACC 204
Qy 85 CysAlaThrThrTyrlsLeuValAsnLeuMetValAlaAspLeuLeuTyrlsValLeu-- 103
Db 205 CGTTCCGCTGTACACCTGAACCTGGACCTGGCAGCTGGCGGACCTGATGATGCTGTTCACTA 264
Qy 104 PropheLeuIleIleThrTyrlsSerLeuAspArgTrpPropheGlyGluLeuLeuCys 123
Db 265 CCCCTACTTATCTATACTACGCCAGAGGGGACCACCTGGCCCTTCGAGACCTCGCTGC 324
Qy 124 LysLeuValHisPheLeuPheTyrlsAsnLeuTyrlsSerlleLeuLeuLeuThrCys 143
Db 325 CGCTTTGTACGCTTCTCTTATGCCAATCTACATGGCAGCATCTGTTCCTCCTACCTGC 384
Qy 144 IleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrlsArg--- 162
Db 385 ATTAGCTTCCAGCGCTACTGGGCATCTGCCACCCCTGGCTTCCTGGCACAAGCGTGA 444
Qy 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValleuGlnLeu 182
Db 445 GGTCCGCGTGTGCTGGGTAGTGTGTGAGTCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 504
Qy 183 LeuProThrLeuAlaPheSerHisThrAspTyrlsAsnGlyGlnMetIleTrpTyrlsAsp 202
Db 505 CTGCCACCGCAGTCTTTGCTGCCACAGCATCCACGCCAACCCGACCTGTGTGTGTACGAC 564
Qy 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrlsGlyIleValleuThrLeuSer 222
Db 565 CTGAGCCCAACCATCTCTGTCTACTCGCTACCTGCGCTATGATGGTATGGCCCTCAGCGTCATC 624
Qy 223 GlyPheLeu--SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProA 242
Db 625 GGCTTCTTGTGCTTCATAGCCTTACTGGCTTGTATTGTTCG----- 668

Qy	242	spGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHisProA	262
Db	669	-----CATGCCCGCGCCTGTGTGCGCCAG	693
Qy	262	spHisProThr-----	265
Db	694	GATGCCCGCAGAGGTCCTGTGGCCAGAGCGGCGCAGAGCGGCTCGTATGGCTGTG	753
Qy	266	-----GlyValTrpProLeuHisProLeuPheCysAlaLeuProTrpYHisSerLeuLeuL	284
Db	754	GTGGTGGCAGCGTCTTTTGCCATCAGCTT--	810
Qy	284	eupProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysG	304
Db	811	TACTTGGCTGTGCGTCCACGCCCGGTCTCTTGGCCCTGTGTGGAGACCTTCGGCTGT	870
Qy	304	1yLeuGlnAspMetGluAlaSerGlyGluCysGluGln-----	316
Db	871	GCTTACAAGGCACCTCGGCCCTTCGCCAGTGTCAACAGTGTCTGGAGCCCATTCCTCTTC	930
Qy	317	-----LeuProGlnProSerProValLeuSerPheLysGlyGlyLysA	331
Db	931	TACTTCACACAACAGAGATTCCGGCGGC-AACCCACGATCTCTTACAGAGGCTCACAG-	988
Qy	331	snArgValArgLeuLeuGlnLysLeuArgGlnAenLysLeuGlyGluHisProAlaGlyA	351
Db	989	-----CCAGTGGCA	998
Qy	351	rgLysArg	353
Db	999	GAGGCAGA	1006

RESULT 8

US-08-442-134A-1
Sequence 1, Application US/08442134A
Patent No. 5596088
GENERAL INFORMATION:
APPLICANT: Boucher, Richard C.
APPLICANT: Weisman, Gary A.
APPLICANT: Turner, John T.
APPLICANT: Harden, Thomas K.
APPLICANT: Parr, Claude E.
APPLICANT: Sullivan, Daniel M.
APPLICANT: Erd, Laura
APPLICANT: Lustig, Kevin D.
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
Null Cells Expressing P2U Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell, Seltzer, Park & Gibson
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: No. 5596088th Carolina
COUNTRY: USA
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,134A
FILING DATE: 16-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-71A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-513-974B-57

Alignment Scores:
Pred. No.: 2,01e-34
Score: 413.00
Percent Similarity: 48.61%
Best Local Similarity: 39.01%
Query Match: 21.33%
DB: 3
Length: 984
Matches: 126
Conservative: 31
Mismatch: 146
Indels: 22
Gaps: 8

33 AspMetAsnThrSerGlnGlnGlnGlyLeu-----CysGlnPheSerGluLys 48
10 GACAATGGCACAGGCCAGGCTCTGGGCTTGGCCACCCACACCTGTGTCTACCGGAGAAC 69
49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerLeuPheLeuLeuGlyLeuProLeu 68
70 TTCAGCAACTGCTGCTGCCACCTGTGTATTCGGGGGTGGGGCTGGCGCTGCCGCTG 129
69 AsnGlyThrValLeuTyrPheSerTyrGlnThrLysArgTyrPheSerCysAlaThrThr 88
130 AACATCTGTGTCATACCCAGATCTGCACGTCGGCGCGGCCCTGACCCGCGACGCGCTG 189
89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu--LeuProPheLeuLeu 107
190 TACACCTAAACCTTGTCTGCTGACCTGTATATGCTCTCTCTCTCTCTCTCTCTCTCTCT 249
108 IleThrTyrSerLeuAspArgTyrProPheGlyGluLeuLeuCysLysLeuValHis 127
250 TACAACTATGCCCAAGGTGATCACTGGCCCTTTGGGACTTGGCTGCGCTGGCTGGCGC 309
128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
310 TTCCTCTTCTATGCAACCTGCACGGCAGCATCTCTTCTCTCACCTGCATCAGCTTCAG 369
148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg--ThrArgArgHis 166
370 CGTACCTGGGCACTGCAACCCCTGGCCCTGGCACAAACGTTGGGGCGCGCGGCT 429
167 AlaTyrLeuGlyThrSerThrThrTyrPheAlaLeuValLeuGlnLeuLeuProThrLeu 186
430 GCCTGGCTAGTGTGTAAACGCTGTGGCTGTGGCCCTGACACCCAGTGGCTGCCACACCC 489
187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGln 206
490 ATCTTCGCTGCCACAGGATCAGCGTAACCGCACTGTCTGTATGACTCAGCCGCT 549
207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226
550 GCCCTGGCCACCCACTATATGCTTATGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTG 609
226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246
610 CCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
246 lylGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
667 GATGCGCC-----GGCAGAGCTTGTGCCCGAGGCGCGCTGGCAGCGCGCGCGCATG 720
266 lylValTyr-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281
721 GCCG-TGGTGTGGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 776
282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
777 GACAGCTTACCTGGCAGTGGGCTGACCGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTG 836
302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSer 321
837 TGCAGCGGCTACAAAAGGACCGCGCGCTTGTGCCAGTGCACACAGCGTGTGGACCCCAT 896
322 ProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuGlnLys 338
897 CCTCTTCTAC--TTCACCCAGAAAGTTCGCGCGCGGCGCACCATGATGAGCTCTCTCAGAA 954
339 Leu 339
955 CTC 957

RESULT 12
US-09-461-436B-57
; Sequence 57, Application US/09461436B
; Patent No. 6538107
; GENERAL INFORMATION:

APPLICANT: Shuji Hinuma
 Yasuaki Ito
 Ryo Fujii
 TITLE OF INVENTION: G Protein Coupled Receptor Protein, Production, And Use Thereof
 NUMBER OF SEQUENCES: 61
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Edwards & Angell, LLP
 STREET: 101 Federal Street
 CITY: BOSTON
 STATE: MA
 COUNTRY: USA
 ZIP: 02209
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/461,436B
 FILING DATE: 14-Dec-1999
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/513,974
 FILING DATE: 14-SEP-1995
 APPLICATION NUMBER: PC7/JP95/01599
 FILING DATE: 10-AUG-1995
 APPLICATION NUMBER: 7-093989
 FILING DATE: 19-APR-1995
 APPLICATION NUMBER: 7-057186
 FILING DATE: 16-MAR-1995
 APPLICATION NUMBER: 7-007177
 FILING DATE: 20-JAN-1995
 APPLICATION NUMBER: 6-326611
 FILING DATE: 28-DEC-1994
 APPLICATION NUMBER: 6-270017
 FILING DATE: 02-NOV-1994
 APPLICATION NUMBER: 6-236357
 FILING DATE: 30-SEP-1994
 APPLICATION NUMBER: 6-236356
 FILING DATE: 30-SEP-1994
 APPLICATION NUMBER: 6-189274
 FILING DATE: 11-AUG-1994
 APPLICATION NUMBER: 6-189273
 FILING DATE: 11-AUG-1994
 APPLICATION NUMBER: 6-189272
 FILING DATE: 11-AUG-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: CONLIN, DAVID G.
 REGISTRATION NUMBER: <Unknown>
 REFERENCE/DOCKET NUMBER: 45753 DIV2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-439-4444
 TELEFAX: 617-439-4170
 INFORMATION FOR SEQ ID NO: 57:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 984 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 57:
 US-09-461-436B-57

Alignment Scores:
 Pred. No.: 2,01e-34 Length: 984
 Score: 413.00 Matches: 126
 Percent Similarity: 48.61% Conservative: 31
 Best Local Similarity: 39.01% Mismatches: 146
 Query Match: 21.33% Indels: 22
 DB: 4 Gaps: 8

US-10-763-972-2 (1-360) x US-09-461-436B-57 (1-984)

33 AspMetAsnThrSerGlnGlnGlyLeu-----CysGlnPheSerGluLys 48
 10 GACAAATGCACAGCCAGGCTCTGGCTTGCACCCACCACTGTGTCTACCGGAGAAC 69
 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 68
 70 TTCAAGCAACTGCTGCTGCCACTGTATTGGCGGTGCTGGCGGCTGGCTGCGCTG 129
 69 AsnGlyThrValLeuTyrHisSerTyrGlyGlnThrLysArgTyrSerCysAlaThrThr 88
 130 AACATCTGTGTATTACCAAGATCTGCAGCTCCCGCGGCGCTGCACCGCAGCGCGCTG 189
 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 107
 190 TACACCTAAACCTTGCTGTGCTGACCTGTATATGCTGTCTCCCTGCCCTGTCTATC 249
 108 IleThrTyrSerLeuAspArgTyrProPheGlyGluLeuLeuCysLysLeuValHis 127
 250 TACAACTATGCCCCAAGGTGATCACTGGCCCTTTGGCGACTTGGCTGCGCTGGTCCGC 309
 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
 310 TTCTCTTCTATGCCAACCTGCACGGCAGCATCTCTTCTCCTCAGCTCATCAGCTTCCAG 369
 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166
 370 CGCTACCTGGGCATCTGCCACCGCTGCCCTTGGCAACCGTGGCGGCGCGCGGCT 429
 167 AlaTyrLeuGlyThrSerThrTyrAlaLeuValLeuLeuGlnLeuLeuProThrLeu 186
 430 GCTGGCTAGTGTGTAACTGCTGGCTGGCTGACAAACCGTGCCTGCCACAGCC 489
 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGln 206
 490 ATCTTCGCTGCCACAGGCATCGAGCGTAACCGCATGTCTGCTATGACCTCAGCCCGCT 549
 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuLeuLeuSerGlyPheLeu--S 226
 550 GCCCTGGCCACCCACTATATGCCCTATGGCATGGCTCTCCTGCTGCTGCTGCTGCTG 609
 226 exLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlnProAspGlnAlaArg 246
 610 CCCTTTGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
 246 IyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
 667 GATGCCCC-----GGCAGAGCCTGTGGCCCGAGAGCGGCTGGCAGCGCGCGCGCTG 720
 266 IyValTyr-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281
 721 GCCG-TGGTGGTGGCTGCTGCTTGGCCATCAGCTT---CCTGCTTTTACATCACCAC 776
 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
 777 GACAGCCTACCTGGCAGTGGGCTCGACCGCGGCGCTCCCTGCTGCTGCTGCTGCTGCT 836
 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGlnCysGluGlnLeuProGlnProSer 321
 837 TGCAGCGCCCTACAAAGCAGCGCGCTTTGGCCAGTGGCAGCGCTGCTGCTGCTGCTGCT 896
 322 ProValLeuSerPheLysGlyLysAsnArgValArg-----LeuLeuGlnLys 338
 897 CCTCTTCTAC--TTCAACCCAGAGAAGTTCCGCGCGGCGGACCATGAGCTCTCTACAGAA 954
 339 Leu 339
 955 CTC 957

RESULT 13

US-08-513-974B-379

; Sequence 379, Application US/08513974B

; Patent No. 6114139

GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ohtaki, Tetsuya
APPLICANT: Fukusumi, Shoji
APPLICANT: Ohgi, Kazuhiro
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
NUMBER OF SEQUENCES: 380
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 379:
SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: CDS
LOCATION: 37...1020
US-08-513-974B-379
Alignment Scores:
Pred. No.: 2,14e-34 Length: 1023
Score: 413.00 Matches: 126
Percent Similarity: 48.61% Conservative: 31
Best Local Similarity: 39.01% Mismatches: 146
Query Match: 21.33% Indels: 22
DB: 3 Gaps: 8
US-10-763-972-2 (1-360) x US-08-513-974B-379 (1-1023)
QY 33 AspMetAsnThrSerGlnGlnGlnGlyLeu-----CysGlnPheSerGlnLys 48
DB 46 GACAATGGCAGCCAGGCTCTGGGTTGCCACCCACCACCTGTCTACCGGAGAAC 105
QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 68
DB 106 TTCAAGCAACTGCTGCTGCCACCTGTGTATTCGGCGGTGTGGCGGTGCTGCCGTG 165
QY 69 AsnGlyThrValLeuTyrPheHisSerTyrGlyGlnThrLysArgTyrSerCysAlaThrThr 88
DB 166 AACATCTGTGTATACCCAGATCTGCAGTCTCCCGCGGCCCTGACCCGACCGCGGTG 225
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 107
DB 226 TACACCTATAACCTTGTCTGTGCTGACCTGTATATGCTCTCTCTCCCTGCCCTGCTCATC 285
QY 108 IleThrTyrSerLeuAspAspArgTyrProPheGlyGlnLeuLeuCysLysLeuValHis 127
DB 286 TACAACTATGCCAAGGTGATCACTGGCCCTTTGGCGACTTCGGCTGCCGCGCGGTG 345
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
DB 346 TTCCTCTTCTATGCCAAGCTGCAGCGGAGCATCTCTCTCTCAGCTGCATCAGTTCCAG 405
QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166
DB 406 CGCTACTCTGGCATCTGCCACCCGCTGGCCCTGGCGACAAACGTTGGGGGCGCGCGGCT 465
QY 167 AlaTyrLeuGlyThrSerThrTyrAlaLeuValValLeuGlnLeuLeuProThrLeu 186
DB 466 GCCTGGCTAGTGTGTGTAAACGTTGGCTGGCGGTGACCAACCCAGTGCCTCCACACGCC 525
QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGln 206
DB 526 ATCTTCGCTGCCACAGGCATCCAGCGTAACCGCATGCTGTGTATGACCTACGCGCGCT 585
QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226
DB 586 GCCCTGGCCACCCACTATATGCCCTATGGCATGGCTCTCAGTGTCTCTGGCGCTG--CCGCTGTGGCGCCAG 702
QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlnProAspGlnAlaArg 246
DB 646 CCCTTGTGCTGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702
QY 246 IyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
DB 703 GATGGCCC-----GGCAGAGCCCTGTGGCCCGAGGCGCGGTGGCAAGCGCGCGCGCATG 756
QY 266 IyValTyr-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281
DB 757 GCCG-TGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 812
QY 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
DB 813 GACAGCCTACCTGGCAGTGGGCTCGACGCGGGCGCTCCCTGCTGCTGCTGCTGCTGCTG 872
QY 302 GlnCysGlyLeuGlnAspMetGlnAlaSerGlyGluCysGluGlnLeuProGlnProSer 321

Db 873 TGCAGCGCCCTACAAAGCAGCGCGCGTTTGGCCAGTGCCAAAGCGGTGCTGGACCCCAT 932
Qy 322 ProValLeuSerPheLysGlyLysAsnArgValArg-----LeuLeuGlnLys 338
Db 933 CCTCTCTTAC--TTACCCAGAGAGAGTTCCGCGCGGAGCCACATGAGCTCTTACAGAA 990
Qy 339 Leu 339
Db 991 CTC 993
RESULT 14
US-09-016-434-1108
; Sequence 1108, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1108:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G1296659
; US-09-016-434-1108
Alignment Scores:
Pred. No.: 6,92e-34 Length: 1571
Score: 411.00 Matches: 126
Percent Similarity: 45.85% Conservative: 34
Best Local Similarity: 36.10% Mismatches: 152
Query Match: 21.23% Indels: 39
DB: 4 Gaps: 9
US-10-763-972-2 (1-360) x US-09-016-434-1108 (1-1571)
Qy 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48
Db 286 GACATGGCAGCGCAGCGGTCTGGGCTTGCCACCACCATGCTGTCTACCGCGAGAAC 345
Qy 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeu 68

Db 346 TTCAGCAACAGCTCTGCTGCCACCTGTGTATTTCGGCGGTGCTGGCGGTGCTGCCGCTG 405
Qy 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88
Db 406 AACATCTGTGTCATTACCCAGATCTGCACGTCCTGGCGCGGCGCTTGCACCGCAGCGCGTG 465
Qy 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTrpValLeu-----LeuProPheIle 107
Db 466 TACACCTTAAACCTTGTCTGTGCTGACCTGTATATGCTCTCTCCCTGCTCTCTCATC 525
Qy 108 IleThrTyrSerLeuAspArgTrpPheGlyGluLeuLeuCysLysLeuValHis 127
Db 526 TACAACTATGCCAAGGTGATCACTGGCCCTTTGGCGACTTTCGCTGCTGCTGCTCGC 585
Qy 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 586 TTCTCTTCTTATGCCAACCTGCACGGCAGCATCTCTTCTCCTCAGCTGCATCAGCTTCCAG 645
Qy 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166
Db 646 CGCTACCTGGGCATCTGCCACCCGCTGGCCCTTGGCACAACCTGGGGCGCGCGGCT 705
Qy 167 AlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuLeuLeuProThrLeu 186
Db 706 GCCTGGCTAGTGTGTAGCGGTGTGGCTGGCTGGCTGACAAACCCAGTGCCTGCCACAGCC 765
Qy 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206
Db 766 ATCTTCGCTGCCACAGGCATCCAGGTAAACCGACTGTCTGTATGACTCAGCCCGCT 825
Qy 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuLeuLeuSerGlyPheLeu--S 226
Db 826 GCCCTGGCCACCCACTATATGCCCTATGGCATGCTCTCCTGCTGCTGCTGCTGCTGCTG 885
Qy 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246
Db 886 CCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 942
Qy 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
Db 943 GATGCCCC-----GGCAGACCTGTGGCCAGCGCGGTGGCAAGCGCGCGCGCATG 996
Qy 266 lyValTrp-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281
Db 997 GCCG-TGGTGGTGGCTGCTGCTGCTTGGCATCAGCTT---CCTGCTTTTTCACATCACA 1052
Qy 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
Db 1053 GACAGCTTACCTGGCAGTGGCTGCGCGCGGGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1112
Qy 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnLeuProGlnProSer 321
Db 1113 TGCAGCGCGCTTACAAAGCAGCGCGCTTGGCCAGTGCACCAACAGCGTGTGGACCCCAT 1172
Qy 322 ProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArgGln 341
Db 1173 CCT-----CTTCTACT-TCACCCAGGA 1195
Qy 342 AsnLysLeuGlyGluHis-----ProAlaGlyArgLysArg 353
Db 1196 AGTTCCGCGCGCAGCACCATGAGCTCTCAGAGAACTCAGAGCAATGCGAGGCGAGG 1255
Qy 354 CysProGlyLeuAsnArgSer 360
Db 1256 GTCGCTAGTCTCTCCAGGTCC 1276
RESULT 15
US-08-459-046-1
; Sequence 1, Application US/08459046
; Patent No. 6008039
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger

APPLICANT: Au-Young, Janice
 APPLICANT: Stuart, Susan G.
 TITLE OF INVENTION: A No. 6008039el Human Purinergic P2U Receptor
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 STREET: 3330 Hillview Avenue
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM: disk
 MEDIUM TYPE: Floppy
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/459,046
 FILING DATE:

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Luther, Barbara J.
 REGISTRATION NUMBER: 33954
 REFERENCE/DOCKET NUMBER: PF-0038 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-852-0195
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 984 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 IMMEDIATE SOURCE:
 LIBRARY: Placenta
 CLONE: 179696
 US-08-459-046-1

Alignment Scores:
 Pred. No.: 3e-33 Length: 984
 Score: 402.00 Matches: 119
 Percent Similarity: 46.60% Conservative: 32
 Best Local Similarity: 36.73% Mismatches: 150
 Query Match: 20.76% Indels: 24
 DE: 3 Gaps: 8

US-10-763-972-2 (1-360) x US-08-459-046-1 (1-984)

QY 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48
 Db 10 GACAAATGGACAGACAGGCTCTGGGCTTGCACCCACACCTGTGTCTACCGGAGAAC 69
 QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 69
 Db 70 TTCAAGCAACTGTCTGCCCTCCACCTGTGTATTGGCGGTGCTGGCGCTGCCCTCCCGCTG 129
 QY 69 AsnGlyThrValLeuThrHisSerTyrGlnThrLysArgTrpSerCysAlaThrThr 88
 Db 130 AACATCTGTGTATACCCAGATCTGCAGTCCCGCGCGCCCTGACCCCGCAGCGCGGTG 189
 QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 107
 Db 190 TACACCTAAACCTTGTCTGCCCTGACCTGCTATATGCTGTCTGCTGCCCTGCTCATC 249
 QY 108 IleThrTyrSerLeuAspAspArgTrpPheGlyGluLeuLeuCysLysLeuValHis 127
 Db 250 TACAACTATGCCCAAGGTGATCATCTGCCCTTTGGCGACTTCGCGCTGCCCGCTGGTCCG 309
 QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
 Db 310 TTCCTCTCTATATGCCAACCTGCACGGGAGGATCTCTTCTCCTCACCTGCATCAGCTTCCAG 369

QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgAsqHis 166
 Db 370 CGCTACCTGGGCATCTGCCACCCCGCTGGCCCCCTGGGCACAAACCTGGGGCGCGCGGCT 429
 QY 167 AlaTrpLeuGlyThrSerThrThrTyrAlaLeuValValLeuGlnLeuLeuProThrLeu 186
 Db 430 CCTTGGCTAGTGTGTAGCGGTGTGGCGGTGACAAACCCAGTGCCTGCCACAGCC 489
 QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206
 Db 490 ATCTTGGCTGCCACAGGCATCCAGCAATACCGCATCTGTGTATGACCTCCGCGCGCT 549
 QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226
 Db 550 GCCCTGGCCACCCACTATATCCCTATGGATGGCTCTCACTGTCATCGGCTTCTGCTG 609
 QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246
 Db 610 CCTTTGTGCTGCTGCTGGCTGCTACTGTCTCTGCGCTG---CCGCTGTGCCCGCAG 666
 QY 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
 Db 667 GATGCCC-----GGCAGAGCCTGTGGCCACAGGCGCGCTGGCAAGCGCGCGCATG 720
 QY 266 ly-----ValTrpProLeuHisProLeuPheCysAlaLeuProTyrH 280
 Db 721 GCCGTGGTGGTGGCTGTCTTTGGCATCAGCTTCTGCTTT-----TCACATC 771
 QY 280 isserLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
 Db 772 ACCAAGACAGGCTACCTGGCAGTGGCTCGACGCGCGCGCTCCCTGCACTGTATTGGAG 831
 QY 300 lySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320
 Db 832 GCCTTGCAGCGGCTCAAAAGGACGCGCGCTTGGCCAGTGCACAGCGGTGCTGGAC 891
 QY 320 roSerProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuG 337
 Db 892 CCCATCTCTTCTAC--TTCACCCAGAAAGATTCCGCGCGGACCCACATGAGCTCCTAC 949
 QY 337 InLysLeu 339
 Db 950 AGAAACTC 957

Search completed: November 10, 2004, 13:18:06
 Job time : 118 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 10, 2004, 12:01:36 ; Search time 625 seconds
(without alignments)
3101.251 Million cell updates/sec

Title: US-10-763-972-2

Perfect score: 1936
Sequence: 1 MSLILLPGRSGSRGAL.....QNKLGHPAGKRCPLNRS 360

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10763972/runat_04112004_093315_8837/app_query.fasta_1.519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10763972@cgn_1_1480 -runat_04112004_093315_8837
-NCPU=6 -ICPU=3 -NO_MMAP=LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1936	100.0	1082	15	US-10-023-586B-1
2	1936	100.0	1082	17	US-10-763-972-1
3	1936	100.0	1334	15	US-10-017-161-707
4	1936	100.0	1334	15	US-10-292-798-619
5	1936	100.0	1343	15	US-10-275-910-1
6	1896	97.9	1108	16	US-10-072-012-165
7	1798	92.9	1002	15	US-10-088-726-25
8	1792	92.6	1017	10	US-09-885-453-3
9	1792	92.6	1017	15	US-10-079-384-5
10	1785	92.2	1020	15	US-10-023-586B-3
11	1785	92.2	1020	17	US-10-763-972-3
12	1785	92.2	1076	16	US-10-333-946-20
13	1541	79.6	850	15	US-10-275-910-4
14	1541	79.6	851	15	US-10-275-910-6
15	921	47.6	681	10	US-09-782-974C-13
16	732	37.8	581	11	US-09-801-944B-47
17	681	35.2	585	11	US-09-801-944B-56
18	445.5	23.1	1651	15	US-10-366-288-41
19	445.5	23.0	1429	10	US-09-077-173A-1
20	445.5	23.0	1429	16	US-10-305-720-1068
21	445.5	23.0	1429	17	US-10-753-695-1
22	442.5	22.9	1098	15	US-10-225-567A-331
23	442	22.8	2025	15	US-10-225-567A-216
24	442	22.8	2025	15	US-10-101-510-722
25	442	22.8	2025	16	US-10-305-720-1482
26	442	22.8	2025	17	US-10-776-827-74
27	440	22.7	2118	15	US-10-101-510-431
28	436	22.5	1922	16	US-10-152-319A-2068
29	420.5	21.7	984	15	US-10-378-087A-41
30	413	21.3	984	15	US-10-378-087A-57
31	411	21.2	1571	15	US-10-354-358-9
32	411	21.2	1571	16	US-10-305-720-1108
33	411	21.2	1832	15	US-10-225-567A-222
34	411	21.2	1832	15	US-10-172-118-994
35	411	21.2	1832	15	US-10-295-027-1071
36	411	21.2	1832	16	US-10-342-887-994
37	402	20.8	984	15	US-10-242-499-1
38	370	19.1	551	10	US-09-791-932-32
39	328	16.2	2424	15	US-10-225-567A-218
40	313.5	16.2	3055	16	US-10-305-720-1456
41	292	15.1	234	11	US-09-801-944B-55
42	281.5	14.5	1020	10	US-09-782-974C-85
43	278.5	14.4	1014	9	US-09-943-798-3
44	278.5	14.4	1014	10	US-09-885-453-2
45	278.5	14.4	1014	14	US-10-023-775B-1

ALIGNMENTS

RESULT 1
US-10-023-586B-1
; Sequence 1, Application US/10023586B
; Publication No. US20030166882A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Ltd. (EP (GB) only)
; APPLICANT: Pfizer Inc. (US, JP, EP except GB)
; APPLICANT: Fidoock, Mark David
; TITLE OF INVENTION: No. US20030166882A1e1 Polypeptide
; FILE REFERENCE: PCT0960AGPR
; CURRENT FILING DATE: 2001-12-18
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,563
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/265,688
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: GB 0101222.8
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 7


```
Qy 101 ValLeuLeuProPheLeuLeuIleThrTyrSerLeuAspArgTyrProPheGlyGlu 120
Db 301 GTGCTATTGCCCTTCTCTATCATCACTACTACTAGATGACAGTGGGCCCTTGGGGAG 360
Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 361 CTGCTCTGAAGCTGGTGCACCTCTCTTATATCAAGCTTTACGGCAGCATCTCTGCTG 420
Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 421 CTGACCTGACATCTCTGTGACCACTTCTTAGGTGTGTGCCACCATGTTGCTGCC 480
Qy 161 TyrArgThrArgHisAlaTyrLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
Db 481 TACGGGACCGCAGGATGCTGCTGGGACACGACACCATGGGCCCTGGTGGTCCCTC 540
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
Db 541 CAGCTGCTGCCACACTGGGCTTCTCCACACGAGTATCATCAATGGCCAGATGATCTGG 600
Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 601 TATGACATGACCGACCCAGAGATTTTGATCGCTTTTGGCTACCGCATAGTTCTGACA 660
Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db 661 TTGCTGCTGCTTCTTCCCTCTCTGCTGCTATTTGGTGTGCTATTCACTGATGGTCAGGAG 720
Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db 721 CCTGATCAAGCCAGAGAGAACTCATGAGACAGGCAACACAGCCGAGCCAGTCCAT 780
Qy 261 ProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrHis 280
Db 781 CCGGACCATCTACTGCTGTGTGGCTCTTCACTCTGTTTGTGCTTCCATATCAC 840
Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
Db 841 TCGCTCTCTTACCTCACCATCTGCTTCTGCTTCTCAGGATGCGCAGCTCTTGATGGC 900
Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
Db 901 AGCCAGTGTGGCTACAGATATGGAGGCTCTGGTGGTGTGAGCAGCTGCCTCAACCC 960
Qy 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
Db 961 AGTCTCTACTTCTTTCAAGGGGGGCAAAATAGACTCAGGCTCCTCCAGAAACTGAGG 1020
Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
Db 1021 CAGAACAGTTGGGTGAGCATCCAGCTGGAGGAGAGATGCCAGGGGTTGAACAGATCT 1080
```

RESULT 3

```
US-10-017-161-707
; Sequence 707, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 707
; LENGTH: 1334
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1334)
FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(715)
FEATURE:
; NAME/KEY: CDS
; LOCATION: (834)..(1134)
US-10-017-161-707
```

Alignment Scores:

```
Pred. No.: 3,18e-211 Length: 1334
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
```

US-10-763-972-2 (1-360) x US-10-017-161-707 (1-1334)

```
Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
Db 117 ATGCTGTCCATTTTGTCTTCCCTTCCAGGGGAAAGCAGAGCGGGAGCGCTCGTGAGCTCTG 176
Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
Db 177 CTCCTGGAGGGAGGCTCCCGGACATGGAGAAGGTGGACATGAATACATCATCAGAAACAA 236
Qy 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db 237 GGTCTCTGCCAGTTCTCAGAGAGATGACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 296
Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTyrGlyGlnThr 80
Db 297 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 356
Qy 81 LysArgTyrSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 357 AAGCGCTGAGCTGGCCACACCTATCTGGTGAACCTGATGGTGGCCGACCTGCTTTAT 416
Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTyrProPheGlyGlu 120
Db 417 GTGCTATTGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCCGGGAG 476
Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 477 CTGCTCTGCAAGCTGGTGCACCTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 536
Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 537 CTGACCTGCATCTCTGTGCACCATCTCTAGGTGTGCCACCCACTGTGTTGCTGCC 596
Qy 161 TyrArgThrArgHisAlaTyrLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
Db 597 TACCGGACCGCAGGACATGCTGGCTGGGCAACAGCACCATCTGGGCCCTGGTGTCTCTC 656
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
Db 657 CAGCTGTCTGCCACACTGGCCCTTCTCCACACGGACTACATCAATGGCCAGATGATCTGG 716
Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 717 TATGACATCACCAGCCAGAGAAATTTGATCGCTTTTTCCTACGGCATAGTTCTGACA 776
Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db 777 TTGCTCTGGCTTTCTTCCCTCTCTGCTATTTGGTGTGCTATTCACTGATGGTCAGGAG 836
Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db 837 CCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGCTCCAT 896
```


QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyHis 280
 Db 897 CCGACCATCTCTACTGGTGTGGCTCTTACCCCTCTGTTTGTGGCTTCCATATCAC 956
 QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 Db 957 TCGCTCTCTACTCTACCATCTGCTTCTGCTTCTCAGGACTGCGCAGCTCTTGATGGC 1016
 QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
 Db 1017 AGCCAGTGTGGCTTACAGATATGAGGAGCTCTGGTGTGAGTGTGAGCAGCTTCAACCC 1076
 QY 321 SerProValLeuSerPheLeuGlyGlyLeuAsnArgValArgLeuGlnLeuLysLeuArg 340
 Db 1077 AGTCTGTACTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTGAGG 1136
 QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
 Db 1137 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAAGAGATGCCACGGTTGAACAGATCT 1196

RESULT 4

US-10-292-798-619
 ; Sequence 619, Application US/10292798
 ; Publication No. US20030235833A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SUWA, MAKIKO
 ; APPLICANT: ASAI, KIYOSHI
 ; APPLICANT: AKIYAMA, YUTAKA
 ; APPLICANT: ABURATANI, HIROYUKI
 ; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
 ; FILE REFERENCE: 084335/166
 ; CURRENT APPLICATION NUMBER: US/10/292,798
 ; PRIOR FILING DATE: 2002-11-13
 ; PRIOR APPLICATION NUMBER: 10/017,161
 ; PRIOR FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: JP 2001-246769
 ; PRIOR FILING DATE: 2001-06-18
 ; NUMBER OF SEQ ID NOS: 2070
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 619
 ; LENGTH: 1334
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; LOCATION: source
 ; FEATURE:
 ; LOCATION: (1)...(1334)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (201)...(715)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (834)...(1134)
 ; US-10-292-798-619

Alignment Scores:
 Pred. No.: 3,18e-211 Length: 1334
 Score: 1936.00 Matches: 360
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 15 Gaps: 0

US-10-763-972-2 (1-360) x US-10-292-798-619 (1-1334)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
 Db 117 ATGCTGTCCATTTTGTCTTCTTCCAGGGAGCAGAGCGGGAGCGGTGTGGAGCTCTG 176
 QY 21 LeuLeuGluClyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGln 40
 Db 177 CTCTGGAGGAGCGCTCCCGGACATGGAGAGGTGGACATGAATACATCACAGGAACAA 236

QY 41 GlyLeuCysGlnPheSerGluLysTyrlsGlnValTyrlsLeuSerLeuAlaTyrlsSerIle 60
 Db 237 GGTCTCTGCGAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGCGCTACAGTATC 296
 QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
 Db 297 AICTTTATCCCTAGGCTGCCACTAAATGGCACTGCTTTGTGGCACTCTCTGGGCCAAACC 356
 QY 81 LysArgTrpSerCysAlaThrThrTyrlsLeuValAsnLeuMetValAlaAspLeuLeuTyrl 100
 Db 357 AAGCGCTGGAGTGTGCCACCACTTATCTGGTGAACCTGATGGTGGCGCACTGCTTTAT 416
 QY 101 ValLeuLeuProPheLeuIleThrTyrlsSerLeuAspAspArgTrpProPheClyGlu 120
 Db 417 GTGCTATTGGCCCTTCTCATCATCACTACTACCTAGTAGACAGTGGCCCTCTGGGGAG 476
 QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrlsAsnLeuTyrlsGlySerIleLeuLeu 140
 Db 477 CTGCTCTGCAAGCTGTGCTGCTTCTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTG 536
 QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 Db 537 CTGACCTGCATCTCTGTGCCACCACTTCTAGGTGTGTGCCACCACTGCTGTCTGCTGCC 596
 QY 161 TyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
 Db 597 TACCGACCCGAGGATGCTGTGGTGGCAACAGCACCACTGTGGGCCCTGTGGTCTCTC 656
 QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrlsAsnGlyGlnMetIleTrp 200
 Db 657 CAGCTGTGTGCCACACAGTGGCTTCTCCACAGCAGTACATCAATGGCCAGCATCTGG 716
 QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrlsGlyIleValLeuThr 220
 Db 717 TATGACATGACAGCCAGAGAAATTTGATCGCTTTTGGCTACGGCATAGTTCTGACA 776
 QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 Db 777 TGTCTGGCTTCTTCTCCCTCTTGGTCAATTTGGTGTGCTATTCTACTGATGGTCAGGAG 836
 QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
 Db 837 CCGTATCAAGCCAGAGGAGAACCTCATGAGACAGGCAACACAGCCCGAGCCAGTCCAT 896
 QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyHis 280
 Db 897 CCGGACCATCTCTACTGGTGTGGCTTCTCACCTCTGTTTGTGGCTTCCATATCAC 956
 QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 Db 957 TCGCTCTCTTACCTCCTACCATCTGCTTCTGCTTCTCAGGACTGCGCAGCTCTTGATGGC 1016
 QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
 Db 1017 AGCCAGTGTGGCTTACAGATATGAGGAGCTCTGGTGTGAGTGTGAGCAGCTTCAACCC 1076
 QY 321 SerProValLeuSerPheLeuGlyGlyLeuAsnArgValArgLeuGlnLeuLysLeuArg 340
 Db 1077 AGTCTGTACTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTGAGG 1136
 QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
 Db 1137 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAAGAGATGCCACGGTTGAACAGATCT 1196

RESULT 5

US-10-275-910-1
 ; Sequence 1, Application US/10275910
 ; Publication No. US20030166142A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ramakrishnan, Shyam
 ; TITLE OF INVENTION: REGULATION OF HUMAN P2Y-LIKE G PROTEIN-COUPLED RECEPTOR
 ; FILE REFERENCE: 4974.00885
 ; CURRENT APPLICATION NUMBER: US/10/275,910

```

; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/203,582
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/269,857
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (520)..(1373)
US-10-275-910-1

Alignment Scores:
Pred. No.: 12e-210 Length: 3143
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-763-972-2 (1-360) x US-10-275-910-1 (1-3143)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
DB 436 ATGCTGTCCATTTTGTCTTCCAGGGAAGCAGAGGAGCGTCTGGAGCTCTG 495
QY 21 LeuLeuGluGlyAlaSerArgMetGluValValAspMetAsnThrSerGlnGluGln 40
DB 496 CTCTGGAGGAGCCCTCCGGGACATGGAGAGTGATGATACATCAGAGNACAA 555
QY 41 GlyLeuCysGlnPheSerGluTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
DB 556 GGTCTCTGCCAGTTCTCAGAGAAGTCAAGCAAGTCTACTCTCCCTGGCCCTACAGTATC 615
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrHisSerTrpGlyGlnThr 80
DB 616 ATCTTTTATCTAGGGTGCACCTAAATGGCAGCTGTCTTGGGACCTCTGGGGCCAAAC 675
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
DB 676 AAGCGTGGAGCTGTGCCACCACTATCTGTGAACCTGATGGTGGCCACCTGCTTTAT 735
QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120
DB 736 GTGCTATTGCCCTTCTCATCATCACTACTAGATGACAGGTGGCCCTTCGGGGAG 795
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
DB 796 CTGCTCTGCAAGCTGGTGGCACTTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 855
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
DB 856 CTGACCTGCATCTCTGTGCCACCACTTCTAGGTGTGGCCACCACTGTGTGGCTGCC 915
QY 161 TyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
DB 916 TACGGACCCGAGCATGCTGCTGGGACCAACCACTGGGCGCTGGTGGTGGTCC 975
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
DB 976 CAGCTGTGCCACACCTGGCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG 1035
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
DB 1036 TATGACATGACCAAGCAAGAAATTTTGTGGCTTTTGTCCCTACGGCATAGTTCTGACA 1095
QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
DB 1096 TTGCTGGCTTTCTTTCCTCTCTGGTCACTTTGGTGTGTCTATCTACTGATGTCAGGAG 1155

```

RESULT 6

```

US-10-072-012-165
; Sequence 165, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shamkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsbrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07

```

```

QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
DB 1156 CCTGATCAAGCCAGGAGGAACTCTATGAGCAGGCAACACACAGCCAGGCTCAT 1215
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
DB 1216 CCGACCATCTCTACTGGTGTGGCCCTTCAACCTCTGTGTGGCCCTTCAATATC 1275
QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
DB 1276 TCGCTCTCTTCTACCTACCATCTGCTTCTCTCAGGACTGCCAGCTCTTGATGC 1335
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
DB 1336 AGCCAGTGTGGCCTACAGATATGAGGCTCTCGTGTGTGTGTGTGTGTGTGTGTGT 1395
QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
DB 1456 CAGAACCAAGTTGGGTGAGCATCCAGCTGGGAGGAAGATGCCCGAGGTTGAACAGATCT 1515

```

21	LeuLeuGluGlyAlaSerArgAspMetGlyLysValAspMetAsnThrSerGlnGluGln	40
72	CTCTCGAGGAGGCGCTCCCGGGACATGGAGAAGTGGACATGAATACATCACAGGACAA	131
41	GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle	60
132	GGTCTCTGCCAGTCTCAGAGAGTACAGCGAGTCTACCTCTCCCTGGCTCAGTATC	191
61	IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr	80
192	ATCTTTATCTTAGGGCTGCCACTAAATGGCACTGTCTTGTTGGCACTCTGGGGCCAAAC	251
81	LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuTyr	100
252	ARGGCTGGAGCTGTGCCACCACTATCTGGTGAACTGATGTGGCCACCTGCTTTAT	311
101	ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGlu	120
312	GTGCTATTGCCCTCTCTCATCATCACCTACTCACTAGATGACAGTGGCCCTTCGGGGAG	371
121	LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu	140
372	CTGTCTGCAGCTGGTGGTCACTTCCTCTTATATCAACCTTTACGGCAGCATCTGTGTG	431
141	LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro	160
432	CTGACCTGCATCTCTGTGCACCACTTCCTAGGTGTGTGCCACCACTGTGTGTGCTGCC	491
161	TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu	180
492	TACCGGACCCGACGACATGCGCTGGGGACCCAGCACCACTCGGGCCCTGTGTGTCTCT	551
181	GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp	200
552	CAGCTGTGCCCAACACTGGCGCTCTCCACACGGACTACATCAATGGCCAGATGATCTGG	611
201	TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr	220
612	TATGACATGACACGCAAGAGAAATTTTGATCGGCTTTTTCCTACGGCATAGTTCTGACA	671
221	LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu	240
672	TGTCTGGCTTCTTCTCCCTTGGTCAATTTTGGTGTGTA-RTCACTGATGGTCAGGAG	730
241	ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis	260

```
QY 89 TyrLeuValAsnLeuMetValAlaSerLeuLeuTyrValLeuLeuProPheLeuLeuIle 108
Db 181 TATCTGGTGAACCTGATGCTGGCCGACCTGCTTTATGTGCTATTGCCCTTCCTCATCATC 240
QY 109 ThrTyrSerLeuAspAspArgTrpPheGlyGluLeuLeuCysIysLeuValHisPhe 128
Db 241 ACCTACTCAGTACAGAGGTGGCCCTTCGGGAGCTCTGCAAGCTGTGCACATTC 300
QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148
Db 301 CTGTTCTATATCAACCTTTACGGCAGCATCCCTGCTGCTGACCTGATCTCTGTCACACAG 360
QY 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTrp 168
Db 361 TTCCTAGGTGTGTGCCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 169 LeuGlyThrSerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 188
Db 421 CTGGGACACAGACACACCTGGGCGCTGGTGGTCTCCAGTGTGCTGCCACATCGGCTTC 480
QY 189 SerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGlnGluAsn 208
Db 481 TCCACACAGGACTACATCAATGCCAGATGATCTGCTATGACATGACACAGCCAGCAAGAAAT 540
QY 209 PheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuLeu 228
Db 541 TTTGATCGGCTTTTTCCTACGGCATAGTTCTGACATGCTGGCTTTCTTTCCCTCCTT 600
QY 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248
Db 601 GGTCAATTTTGGTGTGCTATTCTGATGATGCTCAGAGCCTGATCAAGCCAGAGGAGAACCT 660
QY 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268
Db 661 CATGAGACAGGCAACACAGCCGAGCAGGTCCATCCGACCATCTACTGTGTGTGG 720
QY 269 ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu 288
Db 721 CCTCTTCACCCCTCTGTTTTGGCCCTTCATATACATCGCTCTCTTCTACCTACCATCTG 780
QY 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
Db 781 CTCTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGGCAGCCAGGTGGCCTACAAGATATG 840
QY 309 GluAlaSerGlyGluCysGlnLeuLeuProGlnProSerProValLeuSerPheLeuGly 328
Db 841 GAGGCCCTCTGCTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTCTGATCTTTCTTTCAAGGGG 900
QY 329 GlyLysAsnArgValArgLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisPro 348
Db 901 GGCAAAATAGATCAGCTCTCTCCAGAACTGAGGAGCAACAAGTTGGTGAGCATCCA 960
QY 349 AlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
Db 961 GCTGGGAGGAAGAGATGCCAGGGTTGACAGATCT 996
```

```
RESULT 8
US-09-885-453-3
; Sequence 3, Application US/09885453
; Publication No. US20030088080A1
; GENERAL INFORMATION:
; APPLICANT: Communi, Didier
; TITLE OF INVENTION: RECEPTOR GPCR10
; FILE REFERENCE: 9409/2082
; CURRENT APPLICATION NUMBER: US/09/885,453
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/885,453
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1017
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DNA sequence
; LOCATION: (1)..(1017)
; OTHER INFORMATION: GPCR10 DNA sequence
US-09-885-453-3

Alignment Scores:
Pred. No.: 6,55e-195 Length: 1017
Score: 1792.00 Matches: 338
Percent Similarity: 99.71% Conservativity: 0
Best Local Similarity: 99.71% Mismatches: 1
Query Match: 92.56% Indels: 1
DB: 10 Gaps: 0

US-10-763-972-2 (1-360) x US-09-885-453-3 (1-1017)
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgGlySerArgGlyValAlaLeu 20
Db 1 ATGTGTTCATTTTGTCTTCTTCAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
Db 61 CTCCTGGAGGAGGAGCTCCGGGACATGGAGAGGTGGACATGAATACATACAGAGAAACA 120
QY 41 GlyLeuCysGlnPheSerGlyLysTyrIysGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db 121 GGTCTCTGCCAGTTCTCAGAGAGATACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 180
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThr 80
Db 181 ATCTTATCTAGGCTGCCCTAAATGGCACTGCTCTGTGGCACTCTCTGGGCGCAAAACC 240
QY 81 LysArgTrpSerCysAlaThrThrThrThrValLeuValAsnLeuMetValAlaAspLeuTyr 100
Db 241 AAGCGCTGAGCTGTGCCACCATCTATCTGTGTGAACCTGTATGGTGGCGGAGCTGCTTTAT 300
QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpPheGlyGlu 120
Db 301 GTGCTATTGGCTTCTCATCATCACCTACTCATAGATGACAGTGGCCCTCTGGGGAG 360
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 361 CTGCTCTCAAGCTGTGTGCACTTCTGTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 421 CTGACCTGCACTCTGTGTGCAACGATTCCTAGTGTGTGTCACCCACTCTGTTCGTGCCC 480
QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
Db 481 TACCGGACCGGAGGAGATGCTGCTGGTGGGACACAGACCTGGGCGCTGGTGTCTCTC 540
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
Db 541 CAGCTGTGTCACACACTGCTGCTTCTCCACAGGAGTACATCAATGGCAGATGATCTGG 600
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 601 TATGACATGACCGAGAGAGATTTTGATCGCTTTTGGCTACGGCATAGTCTGTGACA 660
QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db 661 TTGCTCTGGCTTTT-TCCCTCTCTGTGCTATTTGTGTGCTATTCACGTATGTCAGGAG 719
QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db 720 CTGTATCAAGCCAGGAGGAGAACCTCATGAGGACAGGACACAGCCGAGCCAGGTCAT 779
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
Db 780 CCGGACCATCTACTGTGTGTGGCTCTTACCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 839
```

QY 281 SerLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 Db 840 TCGCTCCCTTCACTCACTCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGC 899
 QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
 Db 900 AGCCAGTGTGGCCATACAGATATGAGGCCCTCTGCTGAGTGTGAGCAGCTGCCCTCAACCC 959
 QY 321 SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuGlnLysLeu 339
 Db 960 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGATCAGGCTCTCCAGAACTG 1016

RESULT 9

US-10-079-384-5
 ; Sequence 5, Application US/10079384
 ; Publication No. US20030108986A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Communi, Didier
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS COMPRISING G-PROTEIN COUPLED RECEPTORS
 ; FILE REFERENCE: 9409/2132
 ; CURRENT APPLICATION NUMBER: US/10/079,384
 ; CURRENT FILING DATE: 2002-02-20
 ; PRIOR APPLICATION NUMBER: US 09/885,453
 ; PRIOR FILING DATE: 2001-06-20
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 1017
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1017)
 ; OTHER INFORMATION:
 US-10-079-384-5

Alignment Scores:
 Pred. No.: 6.55e-195 Length: 1017
 Score: 1792.00 Matches: 338
 Percent Similarity: 99.71% Conservative: 0
 Best Local Similarity: 99.71% Mismatches: 1
 Query Match: 92.56% Indels: 1
 DB: 15 Gaps: 0

US-10-763-972-2 (1-360) x US-10-079-384-5 (1-1017)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
 Db 1 ATGCTGTCCATTTCCTTCTCAGGGGAAGCAGAGCGCGAGCGCTGTGGAGCTCTG 60
 QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
 Db 61 CTCTGTGAGGGAGCTCCCGGACATCGAGAGGTGGACATGATATACATCACAGGACAA 120
 QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerIleValTyrSerIle 60
 Db 121 GGTCTGTCCAGATTTCTCAGAGAGATACAGCAAGTCTTACCTCTCCCTGGCCCTACAGTATC 180
 QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
 Db 181 ATCTTATCTAGGCTGCCACTAAATGGACATGCTTGTGGACCTCTCCGGGCCCAACCC 240
 QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
 Db 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGAACCTGATGTGGCCGACCTGCTTTAT 300
 QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120
 Db 301 GTGCTATGGCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 360
 QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
 Db 361 CTGCTCTGCAAGCTGGTGCACTTCCITGTTCTATATCAACCTTTACGGCAGCATCTGCTG 420

QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 Db 421 CTGACTGCATCTCTGTGCACCAAGATTTTCATGGCTTTTGGCTACGGCATAGTTCTGACA 480
 QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeu 180
 Db 481 TACCGAGCCCGCAGGCATGCTGCTGGCGCACACACACACCTGGCCCTCGTGGTCTC 540
 QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
 Db 541 CAGCTGTGCTGCCACACTGGCTTCTCCACACAGGACTACATCATGGCCAGATGATCTGG 600
 QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
 Db 601 TATGACATGACCCAGCAGAGAAATTTTCATGGCTTTTGGCTACGGCATAGTTCTGACA 660
 QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 Db 661 TTGCTGGCTTTT-TCCCTCTCTGTCATTTGTGTGCTTATTCACCTGATGGTCAGGAG 719
 QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
 Db 720 CCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGTCCAT 779
 QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
 Db 780 CCGGACCATCTACTGTGTGGCTTTCACCTCTGTTTGTGGCTTCCATATACAC 839
 QY 281 SerLeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 Db 840 TCGCTCTCTTACCTCACCATCTGCTTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGC 899
 QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
 Db 900 AGCCAGTGTGGCTACAGATATGAGGCCCTCTGCTGAGTGTGAGCAGCTGCCCTCAACCC 959
 QY 321 SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuGlnLysLeu 339
 Db 960 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGATCAGGCTCTCCAGAACTG 1016

RESULT 10

US-10-023-586B-3
 ; Sequence 3, Application US/10023586B
 ; Publication No. US20030166882A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pfizer Ltd. (EP (GB) only)
 ; APPLICANT: Pfizer Inc. (US, JP, EP except GB)
 ; APPLICANT: Fidock, Mark David
 ; TITLE OF INVENTION: No. US20030166882A1el Polypeptide
 ; FILE REFERENCE: PCI0960AGR
 ; CURRENT APPLICATION NUMBER: US/10/023,586B
 ; CURRENT FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: GB 0030855.1
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: US 60/260,563
 ; PRIOR FILING DATE: 2001-01-09
 ; PRIOR APPLICATION NUMBER: US 60/265,688
 ; PRIOR FILING DATE: 2001-02-01
 ; PRIOR APPLICATION NUMBER: GB 0101222.8
 ; PRIOR FILING DATE: 2001-01-17
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1020
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-023-586B-3

Alignment Scores:

Pred. No.: 4.17e-194 Length: 1020
 Score: 1785.00 Matches: 337
 Percent Similarity: 99.12% Conservative: 1

Best Local Similarity: 98.83%		Matches: 1	
Query Match: 92.20%		Indels: 2	
DB: 15		Gaps: 0	
US-10-763-972-2 (1-360) x US-10-023-586B-3 (1-1020)			
QY	1	MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu	20
DB	1	ATGCTGTCCATTTTCTCTCCAGGGAAGCAGAGGGAGCGCTGGAGCTCTG	60
QY	21	LeuLeuGluGlyAlaSerArgAspMetGluLeuValAspMetAsnThrSerGlnGluGln	40
DB	61	CTCCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATGAATACATCAGAGAACAA	120
QY	41	GlyLeuCysGlnPheSerGluLeuValTyrGlnValTyrLeuSerLeuAlaTyrSerIle	60
DB	121	GGTCTCTGCAGTCTTCAGAGAGTACACAGAGTCTACCTCTCCCTGGCCCTACAGTATC	180
QY	61	IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThr	80
DB	181	ATCTTTATCTAGGCTGCGACCTAAATGGCACTGTCTTGGCACTCTCTGGGCGCAACCC	240
QY	81	LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr	100
DB	241	AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGTGGTGGCGGACCTGCTTTAT	300
QY	101	ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpPropheGlyGlu	120
DB	301	GTGCTATTGGCCCTTCTCATCATCACTACCTACATAGATGACAGGTGGCCCTTCGGGAG	360
QY	121	LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu	140
DB	361	CTGCTCTGCAAGCTGTGCACCTTCTCTATATCAACCTTTACGGGACGATCTCTGCTG	420
QY	141	LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro	160
DB	421	CTGACCTGATCTCTGTGGACCACTCTAGGTGTGGCACCACTGTGTCTGCTGCC	480
QY	161	TyrArgThrArgGHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeu	180
DB	481	TACCGGACCGCAGGAGATGCTGGTGGGACACGACCACTGGGCGCTCTGGTGTCTC	540
QY	181	GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp	200
DB	541	CAGTGTGCCACACTGGGCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG	600
QY	201	TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr	220
DB	601	TATGACATGACCCAGAGAGATTTTGTATCGGCTTTTTCCTACGCGCATAGTCTGACA	660
QY	221	LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG	240
DB	661	TTGTCTGGCTTCTTTTCCCCCTCTCTGGTCAATTTTGTGTGCTATTCTAGTATGTCAGG	720
QY	240	luProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH	260
DB	721	AGCCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCGTCC	780
QY	260	isProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrH	280
DB	781	ATCCGGACCATCTACTGTGTGGGCTCTTACCCCTCTGTGTGTGGCCCTTCCATATC	840
QY	280	isSerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG	300
DB	841	ACTGCTCTCTTACTTACCATCTGCTTTCTGCTTCTCAGGACTGCCAGCTCTTGATG	900
QY	300	lySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP	320
DB	901	GCACCCAGTGTGGCTCAAGATATGAGGCGCTCTGGTGTAGTGTGAGCAGCTGCGCTCAAC	960
QY	320	roSerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeu	339
DB	961	CCAGTCTGTACTTCTTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCTCCAGAAACTG	1019

RESULT 11			
US-10-763-972-3			
; Sequence 3, Application US/10763972			
; Publication No. US20040137500A1			
; GENERAL INFORMATION:			
; APPLICANT: PFIZER INC.			
; APPLICANT: FIDOCK, MARK DAVID			
; TITLE OF INVENTION: NOVEL POLYPEPTIDE			
; FILE REFERENCE: PC10960B			
; CURRENT APPLICATION NUMBER: US/10/763,972			
; CURRENT FILING DATE: 2004-01-23			
; PRIOR APPLICATION NUMBER: GB 0030855.1			
; PRIOR FILING DATE: 2000-12-18			
; PRIOR APPLICATION NUMBER: US 60/260,563			
; PRIOR FILING DATE: 2001-01-09			
; PRIOR APPLICATION NUMBER: US 60/265,688			
; PRIOR FILING DATE: 2001-02-01			
; PRIOR APPLICATION NUMBER: GB 0101222.8			
; PRIOR FILING DATE: 2001-01-17			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 3			
; LENGTH: 1020			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-763-972-3			
Alignment Scores:			
Pred. No.:	4.17e-194	Length:	1020
Score:	1785.00	Matches:	337
Percent Similarity:	99.12%	Conservative:	1
Best Local Similarity:	98.83%	Mismatches:	1
Query Match:	92.20%	Indels:	2
DB:	17	Gaps:	0
US-10-763-972-2 (1-360) x US-10-763-972-3 (1-1020)			
QY	1	MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu	20
DB	1	ATGCTGTCCATTTTGTCTCTCCAGGGAAGCAGAGGGAGCGCTGGAGGACTCTG	60
QY	21	LeuLeuGluGlyAlaSerArgAspMetGluLeuValAspMetAsnThrSerGlnGluGln	40
DB	61	CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCAGAGAACAA	120
QY	41	GlyLeuCysGlnPheSerGluLeuValTyrGlnValTyrLeuSerLeuAlaTyrSerIle	60
DB	121	GGTCTCTGCAGTCTTCAGAGAGTACACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC	180
QY	61	IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThr	80
DB	181	ATCTTTATCTAGGCTGCGACCTAAATGGCACTGTCTTGGCACTCTCTGGGCGCAACCC	240
QY	81	LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr	100
DB	241	AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGTGGTGGCGGACCTGCTTTAT	300
QY	101	ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpPropheGlyGlu	120
DB	301	GTGCTATTGGCCCTTCTCATCATCACTACCTACATAGATGACAGGTGGCCCTTCGGGAG	360
QY	121	LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu	140
DB	361	CTGCTCTGCAAGCTGTGCACCTTCTCTATATCAACCTTTACGGGACGATCTCTGCTG	420
QY	141	LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro	160
DB	421	CTGACCTGATCTCTGTGGACCACTCTAGGTGTGGCACCACTGTGTCTGCTGCC	480
QY	161	TyrArgThrArgGHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeu	180
DB	481	TACCGGACCGCAGGAGATGCTGGTGGGACACGACCACTGGGCGCTCTGGTGTCTC	540
QY	181	GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp	200
DB	541	CAGTGTGCCACACTGGGCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG	600
QY	201	TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr	220
DB	601	TATGACATGACCCAGAGAGATTTTGTATCGGCTTTTTCCTACGCGCATAGTCTGACA	660
QY	221	LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG	240
DB	661	TTGTCTGGCTTCTTTTCCCCCTCTCTGGTCAATTTTGTGTGCTATTCTAGTATGTCAGG	720
QY	240	luProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH	260
DB	721	AGCCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCGTCC	780
QY	260	isProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrH	280
DB	781	ATCCGGACCATCTACTGTGTGGGCTCTTACCCCTCTGTGTGTGGCCCTTCCATATC	840
QY	280	isSerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG	300
DB	841	ACTGCTCTCTTACTTACCATCTGCTTTCTGCTTCTCAGGACTGCCAGCTCTTGATG	900
QY	300	lySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP	320
DB	901	GCACCCAGTGTGGCTCAAGATATGAGGCGCTCTGGTGTAGTGTGAGCAGCTGCGCTCAAC	960
QY	320	roSerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeu	339
DB	961	CCAGTCTGTACTTCTTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCTCCAGAAACTG	1019


```

QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
Db 541 CAGCTGCTGCCACACTGGCTTCTCCACACGAGCTACATCATGGCCAGATGATCTGG 600
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 601 TATGACATGACCCAGACAGAAATTTGATCGGCTTTTGCTACGCAATGTTCTGACA 660
QY 221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG 240
Db 661 TTGCTGGCTTCTTCCCTCTCTGGTCAATTTGGTGCTATTACCTGATGGTCAGG 720
QY 240 LuProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH 260
Db 721 AGCCTGATCAAGCCAGAGAGACCTCATGAGGACAGCAACACAGCCCGAGCCAGGTCC 780
QY 260 isProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrH 280
Db 781 ATCCGACCATCCACTGCTGTGGCTCTTACCTCTGTTTGGCTTCTCCATATC 840
QY 280 isSerLeuLeuLeuProHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
Db 841 ACTCGCTCTTCTACCTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATG 900
QY 300 lysSerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320
Db 901 GCACCCAGTGTGGCTTCAAGATATGAGGCTCTGGTGTGAGTGAGCAGCTGCCTCAAC 960
QY 320 roSerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuGlnLysLeu 339
Db 961 CCAGTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCCAGAACTG 1019

RESULT 12
US-10-333-946-20
; Sequence 20, Application US/10333946
; Publication No. US20040023252A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.
; APPLICANT: ARVIZU, Chandra S.; LAL, Preeti G.
; APPLICANT: BURFORD, Neil; YUE, Henry
; APPLICANT: GANDHI, Ameena R.; ELLIOTT, Vicki S.
; APPLICANT: RANKUNAR, Jayalaxmi; BAUGHN, Mariah R.
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Navinder K.
; APPLICANT: HAFALIA, April J.A.; VAO, Monique G.
; APPLICANT: LU, Yan; TRIBOULEY, Catherine M.
; APPLICANT: POLICKY, Jennifer L.; KEARNEY, Liam
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.
; APPLICANT: LEE, Ernestine A.; DING, Li
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0176 USN
; CURRENT APPLICATION NUMBER: US/10/333,946
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/23433
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/221,478
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/223,268
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/227,054
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 60/231,121
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,243
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 60/232,691
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/235,146
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1076

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7474806CB1
US-10-333-946-20

Alignment Scores:
Pred. No.: 4,53e-194 Length: 1076
Score: 1785.00 Matches: 337
Percent Similarity: 99.12% Conservatives: 1
Best Local Similarity: 99.83% Mismatches: 1
Query Match: 92.20% Indels: 2
DB: 16 Gaps: 0

US-10-763-972-2 (1-360) x US-10-333-946-20 (1-1076)

```

```

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
Db 57 ARGCTGTCCATTTTCTCTTCCAGGGAAGACAGCGGAGCCGTCTGGAGCTCTG 116
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGln 40
Db 117 CTCCTGGAGGAGGCGCTCCCGGACATGGAGAAGGTGGACATGAATACATCACAGGAACA 176
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db 177 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCCTCAGTATC 236
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTyrGlyGlnThr 80
Db 237 ATCTTTATCTTAGGGCTGCCATTAATGGCACTGTCTTGTGGCACTCTGGGGCCAAACC 295
QY 81 LysArgTyrSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 297 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCGCCAGCTGCTTAT 356
QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTyrProPheGlyGlu 120
Db 357 GTGCTATTGGCCCTTCTCTCATCATCCTACTCTAGTAGACAGTGGCCCTTCGGGAG 416
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 417 CTGCTCTCAAGCTGGTGCACTTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 476
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 477 CTGACCTGCATCTCTGTGCACCACTTCTAGTGTGTGGCACTGCTGCTGCTGCC 536
QY 161 TyrArgThrArgArgHisAlaTyrLeuGlyThrSerThrThrTyrAlaLeuValValLeu 180
Db 537 TACCGGACCCGAGCATGCTGTGGCCACAGCACCACCTGGGCCCTGGTGGTCTC 596
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
Db 597 CAGCTGTGCCCACTGCGCTTCTCCACACGAGCTACATCATATGGCCAGATGATCTGG 656
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 657 TATGACATGACCCAGACAGAAATTTGATCGGCTTTTGGCTACGCAATAGTCTGACA 716
QY 221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG 240
Db 717 TTGCTGGCTTCTTCTCCCTCTCTGGTCAATTTGGTGCTATTACCTGATGGTCTCAGG 776
QY 240 LuProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH 260
Db 777 AGCCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGCAACACAGCCCGAGCCAGGTCC 836
QY 260 isProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrH 280
Db 837 ATCCGACCATCTCTACTGCTGTGGCTCTTCCACCTCTTGTGGCTTCTCCATATC 896

```


QY 280 iSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
Db 897 ACTCGCTCCTTCTAACCCTACCATCTGCTTTCTGCTTCTCAGACTGCCAGCTCTGATG 956
QY 300 lSerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnGlnLeuProGlnP 320
Db 957 GCACCCAGTGTGGCCTTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 1016
QY 320 roSerProValLeuSerPheLysGlyGlyAsnArgValArgLeuLeuGlnLysLeu 339
Db 1017 CCAGTCTCTGACTTCTTTCAAGGGGGCAAAATAGATCAGGCTCTCCAGAAACTG 1075
RESULT 13
US-10-275-910-4
; Sequence 4, Application US/10275910
; Publication No. US20030166142A1
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Shyam
; TITLE OF INVENTION: REGULATION OF HUMAN P2Y-LIKE G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCES: 4974.00885
; CURRENT APPLICATION NUMBER: US/10/275,910
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/203,582
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/269,857
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-275-910-4
Alignment Scores:
Pred. No.: 2,81e-166 Length: 850
Score: 1541.00 Matches: 283
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.60% Indels: 0
Gaps: 15
US-10-763-972-2 (1-360) x US-10-275-910-4 (1-850)
QY 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeuCysGlnPheSerGluLys 48
Db 1 ATGGAGAGGTGGACATGAATACATACAGAAACAGGCTCTGCCAGTCTCCAGAGAAG 60
QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 68
Db 61 TACAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTTTATCTAGGCTGCCACTA 120
QY 69 AsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88
Db 121 AATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACACC 180
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeuIleIle 108
Db 181 TAUCTGTGTAACCTGATGATGGTGGCGGACCTGCTTTATGCTATGCTATGCCCTTCCTCATC 240
QY 109 ThrTyrSerLeuAspAspA-gftrpProPheGlyGluLeuLeuCysLysLeuValHisPhe 128
Db 241 ACCTACTCCTAGATGACAGTGGCCCTCTGGGAGAGCTGCTCTGCAAGCTGGTGCATTC 300
QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148
Db 301 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGTGACCTGCTATCTGTGGCACCAG 360
QY 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgA-gHisAlaTrp 168
Db 361 TTCCTAGTGTGGCCACCACTGCTGTTGGCTGCCCTACCGACCCGAGCATGCGCTGG 420
QY 169 LeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeuLeuProThrLeuAlaPhe 198

Db 421 CTGGGCACACAGCACCTGGGCCCTGGTGGTCTCTCCAGCTGTGCCACACTGGGCTTC 480
QY 189 SerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGluAsn 208
Db 481 TCCACACAGGACTACATCAATGGCCAGATGATCTGGTATGACATGACACCCAGAGAAT 540
QY 209 PheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuLeu 228
Db 541 TTTGATCGGCTTTTGGCTACGGCATAGTCTTGCATTTGTGCTTCTTCTTCCCTCCTT 600
QY 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGlnProAspGlnAlaArgGlyGluPro 248
Db 601 GGTCAATTTGGTGTGCTATTCACTGATGGTCAGGAGCTGATCAAGCCAGAGGAGAACCT 660
QY 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268
Db 661 CATGAGACAGGCAACACAGCCGAGCAGTCCATCCGACCATCTCTACTGTGTGTGG 720
QY 269 ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu 288
Db 721 CCTCTTACCCTCTGTTTGTGGCTTCCATATCACTCGTCTCTTCTTACCTCACCATCTG 780
QY 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
Db 781 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCTTACAGATATG 840
QY 309 GluAlaSer 311
Db 841 GAGGCCTCT 849
RESULT 14
US-10-275-910-6
; Sequence 6, Application US/10275910
; Publication No. US20030166142A1
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Shyam
; TITLE OF INVENTION: REGULATION OF HUMAN P2Y-LIKE G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCES: 4974.00885
; CURRENT APPLICATION NUMBER: US/10/275,910
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/203,582
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/269,857
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-275-910-6
Alignment Scores:
Pred. No.: 2,82e-166 Length: 851
Score: 1541.00 Matches: 283
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 79.60% Indels: 0
Gaps: 15
US-10-763-972-2 (1-360) x US-10-275-910-6 (1-851)
QY 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeuCysGlnPheSerGluLys 48
Db 1 ATGGAGAGGTGGACATGAATACATACAGGAAACAGGCTCTGCCAGTCTTCAGAGAAG 60
QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 68
Db 61 TACAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTTTATCTAGGCTGCCACTA 120
QY 69 AsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88

Db 121 AATGGCACTGCTTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCACC 180
Qy TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeuLeuLeu 108
Db 181 TATCTGGTGAACCTGATGGTGGCGACTGCTTTATGTGCTATTGCGCTTCTCTCATCATC 240
Qy ThrTyrSerLeuAspAspArgTyrProPheGlyGluLeuLeuCysLysLeuValHisPhe 128
Db 241 ACCTACTCACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTC 300
Qy LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148
Db 301 CTGTTCTATATCAACCTTTACGGCAGATCTCTGCTGACCTGCACTCTCTGGCACCAG 360
Qy PheLeuGlyValCysHisProLeuLeuLeuProTyrArgThrArgArgHisAlaTyr 168
Db 361 TTCCTAGGTGTGCCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy LeuGlyThrSerThrThrTyrAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188
Db 421 CTGGGCAACCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy SerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGlnGluAsn 208
Db 481 TCCACACGGACTACATCATGCGCAGATGATCTGCTGATGACATGACACCGCAGAGAAAT 540
Qy PheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuLeu 228
Db 541 TTTGATCGGCTTTTGTGCTAGCGCATAGTTCTGCATTTGTCTGCTGCTGCTGCTGCT 600
Qy GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248
Db 601 GGTCACTTTGGTGGCTATTCATCATGATGTCAGAGCTGATCAAGCCAGAGAGAACCT 660
Qy HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTyr 268
Db 661 CATGAGGACAGGCAACACAGCCGAGCCAGGTCCATCCGACCATCTACTGCTGTGTG 720
Qy ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu 288
Db 721 CCTCTTCACCTCTGTTTGTGCTTCCATATCATCTGCTGCTGCTGCTGCTGCTGCTG 780
Qy LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
Db 781 CTCTCTCTTCACGAGACTGCCAGCTCTTGTATGGCAGCCAGTGTGGCTACAGATATG 840
Qy GluAlaSer 311
Db 841 GAGGCTCTCT 849

RESULT 15

US-09-782-974C-13
; Sequence 13, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
; FILE REFERENCE: 411USPHRM311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19

; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-974C-13
Alignment Scores:
Pred. No.: 1.35e-95 Length: 510
Score: 921.00 Matches: 170
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 47.57% Indels: 0
Gaps: 10
DB: 10
US-10-763-972-2 (1-360) x US-09-782-974C-13 (1-510)
Qy 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu 102
Db 1 TGGAGCTGTGGCCACCACTATCTGTGTGAACCTGATGTGGCCGACCTGCTTATGTGCTA 60
Qy 103 LeuProPheLeuIleIleThrTyrSerLeuAspAspArgThrProPheGlyGluLeuLeu 122
Db 61 TTGCGCTCTCTCATCATCACTACTCTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 120
Qy 123 CysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThr 142
Db 121 TGCAAGCTGGTGGCACTTCTCTTCTATATCAACCTTTACGCGACGATCTCTGCTGCTG 180
Qy 143 CysIleSerValHisGlnPheLeuGlyValCysHisProLeuLeuCysSerLeuProTyrArg 162
Db 181 TGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGCTGCTGCTGCTGCTGCTG 240
Qy 163 ThrArgArgHisAlaTyrLeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeu 182
Db 241 ACCGCGAGGCAATGCTGCTGGCCACCACTGCGGCTGGTGTGCTGCTGCTGCTGCTGCTG 300
Qy 183 LeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAsp 202
Db 301 CTGCGCCACACTGGCGCTTCTCCACACCGACTACATCAATGGCCAGATGATCTGGTATGAC 360
Qy 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSer 222
Db 361 ATGACACGCAAGAGAAATTTGATCGGCTTTTGGCTTACGCGATAGTCTGACATTTGCT 420
Qy 223 GlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAsp 242
Db 421 GGCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480
Qy 243 GlnAlaArgGlyGluProHisGluAspArg 252
Db 481 CAAGCCAGAGGAGAACCTCATGAGGACAGG 510

Search completed: November 10, 2004, 14:57:07
Job time : 632 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 10, 2004, 10:44:11 ; Search time 3860 Seconds
(without alignments)
3398.524 Million cell updates/sec

Title: US-10-763-972-2
Perfect score: 1936
Sequence: 1 MSLILLPSRSGSRGAL.....QNKLGHPAGKRCPCPLNRS 360

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cg21/USPfo.spool/US10763372/runat_04112004_093314_8797/app.query.fasta_1.519
-DB=EST -QWIT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US10763372 @CGN 1 1 3437 @runat_04112004_093314_8797 -NCPU=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	460.5	23.8	4439	3 AK076364	AK076364 Mus muscu
2	455	23.5	693	6 CD217676	CD217676 pgrin.pk0
3	420.5	21.7	987	9 AY400799	AY400799 Mus muscu
4	416	21.5	987	9 AY400798	AY400798 Pan trogl
5	414.5	21.4	1425	3 CR624871	CR624871 full-leng
6	414.5	21.4	2542	3 AK017378	AK017378 Mus muscu
7	414.5	21.4	3001	3 AK005013	AK005013 Mus muscu
8	411	21.2	987	9 AY400797	AY400797 Homo sapi
9	411	21.2	1515	3 CR612681	CR612681 full-leng

10	411	21.2	1532	3	CR618945	full-leng
11	411	21.2	1564	3	CR626266	full-leng
12	411	21.2	1589	3	CR626754	full-leng
13	411	21.2	1592	3	CR605888	full-leng
14	411	21.2	1594	3	CR606834	full-leng
15	411	21.2	1638	3	CR617080	full-leng
16	411	21.2	1651	3	CR597500	full-leng
17	411	21.2	1686	3	CR626706	full-leng
18	411	21.2	1690	3	CR598765	full-leng
19	398	20.6	845	5	EX740173	EX740173
20	397.5	20.5	1035	5	EX739990	EX379990
21	396.5	20.5	1042	5	EX397649	EX397649
22	387	20.0	639	6	CA352828	624212 NC
23	386	19.9	925	5	BX704227	BX704227
24	380	19.6	828	6	CD110272	AGENCOURT
25	380	19.6	970	5	EX385626	EX385626
26	380	19.6	993	5	EX394202	EX394202
27	380	19.6	1013	3	CR614658	full-leng
28	380	19.6	1065	5	EX337204	EX337204
29	380	19.6	1090	3	CR598708	full-leng
30	378	19.5	851	4	BI833118	603090834
31	370	19.1	955	5	EX354044	EX354044
32	361.5	18.7	972	5	EX337012	EX337012
33	360.5	18.6	641	5	BQ396255	NISC ng19
34	359.5	18.6	627	1	AL675845	AL675845
35	357	18.4	606	8	AZ953874	2M0219117
36	355.5	18.4	730	7	CK359970	AGENCOURT
37	355.5	18.4	1063	4	BM918491	BM918491
38	354	18.3	804	7	BP688283	BP688283
39	353.5	18.3	1054	5	EX364429	EX364429
40	351.5	18.2	1815	7	CN210812	RJA041C08
41	350	18.1	916	5	EX396836	EX396836
42	348.5	18.0	1014	5	EX381745	EX381745
43	345.5	17.8	862	7	CN839909	AGENCOURT
44	341	17.6	892	1	AL525099	AL525099
45	337.5	17.4	898	4	BI819396	603034571

ALIGNMENTS

RESULT 1
LOCUS AK076364
DEFINITION Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732468D08 product:P2Y4 RECEPTOR, full insert sequence.
ACCESSION AK076364
VERSION AK076364.1 GI:26345325
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 11042159
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Db	127	CTACTGGGGCTGCCACTCAATGCGGTGTGTCATTTGGGCGAGATCTGGCTGGCGCCGCAAGCGC	180
Qy	83	TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu	102
Db	187	TTGACGGCGACCAACCATCTACATGTAACTGGCCATGGCCGACCTGCTTATGCTTCG	246
Qy	103	--LeuProPheLeuLeuLeuThrTyrSerLeuAspAspArgTrpProPheGlyGluLeu	121
Db	247	TCCTCCCTCTCTCATCTACAACACTACACCGAAGGATTACTGGCCCTTTGGGGACTTC	306
Qy	122	LeuCysLysLeuValHisPheLeuPheTyrTyrIleAsnLeuTyrGlySerIleLeuLeu	141
Db	307	ACCTGCAAAATCGTCCGCTTCAGATTCTACCAACCTGCACGCGAGCATCCTCTTCCTC	366
Qy	142	ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeu---	160
Db	367	ACCTGCATCAGCTCCAGCGCTACATGGGATCTGCCACCCCTTGGCTCGTGCGACAAA	426
Qy	161	TyArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu	180
Db	427	AAGAAGGAAAGAAAGCTGACGTGGCTGGTGTGCTGCTGGCTGGTGTGCTCATCGTCATC	486
Qy	181	GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp	200
Db	487	CAGTGCCTCGCCACCTTTGCTTCTTCGCTCCACCGGACGAGAGATCGCACTGCTGC	546
Qy	201	TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyValLeuThr	220
Db	547	TATGACCTGAGCCCCCGGACCGCTCCACATCTTCTCCCTATGGCATCAGCTGACC	606
Qy	221	LeuSerGlyPheLeu--SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG	240
Db	607	ATCAGCGCTTCCTGCTGCGCCCTTCGAGCCATCTGGCTGCTGCTACTGCGAGATGCCCGC	666
Qy	240	luPro 241	
Db	667	ATCCT 671	
RESULT 3			
AY400799			
LOCUS			
DEFINITION	AY400799	987 bp DNA linear	GSS 15-DEC-2000
ACCESSION	AY400799	Mus musculus p2RY6 gene, VIRTUAL TRANSCRIPT, partial sequence,	
VERSION	AY400799.1	genomic survey sequence.	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

```

/db_xref="taxon:10090"
<1..>987
/gene="P2RY6"
/locus_tag="HCM0677"

ORIGIN
Alignment Scores:
Pred. No.: 2,97e-30 Length: 987
Score: 420.50 Matches: 117
Percent Similarity: 45.05% Conservative: 47
Best Local Similarity: 32.14% Mismatches: 125
Query Match: 21.72% Indels: 76
DB: 9 Gaps: 9

US-10-763-972-2 (1-360) x AY400799 (1-987)

QY 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeu-----CysGln 44
D 1 ATGGAGCAG--GACATGGCACCATCCAGGCTCCAGGCTTGGCGCCACCACCTGGTC 57
QY 45 PheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerLlellePheLeu 64
D 58 TACCGTGAGGATTTCAACGCGACTCTGCTAACCCGGTATCTCGGTGGTGGTGC 117
QY 65 GlyLeuProLeuAsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArgTrpSer 84
D 118 GGCCTGCCACTGACATCTGGCTGATTCGCCAGATCTGGCATCCCGCGGACCTGACC 177
QY 85 CysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---Leu 103
D 178 CGTTCCGCTGTGTACACCTGACCTGCGCACTGGCGGACCTGATGTATGCTGTCACTA 237
QY 104 ProPheLeullelleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCys 123
D 238 CCCCTACTTATCTAATACTACGCGAGGAGGACCACTGGCCCTTCGGAGACCTCGCGCTGC 297
QY 124 LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerLlelleLeuLeuThrCys 143
D 298 CGCTTTGTACGCTTCTCTTCTATGCAATCTACATGGCAGCATCTGTTCTCTCACTGC 357
QY 144 IleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg--- 162
D 358 ATTAGCTTCACGCGCTACCTGGGATCTGCCACCCCTGCTTCCTGGCACAAGCGTGA 417
QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuLeu 182
D 418 GGTGCGCGTGTCTGTGGTAGTGTGGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 477
QY 183 LeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetLleTrpTyrAsp 202
D 478 CTGCCCCACGCGATCTTGTGTCACAGGATCCAGCGCAACGCGCACTGTGTCTACGAC 537
QY 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSer 222
D 538 CTGAGCCCCATCCTGTCTACTCGTACTGCTTATGCTATGATGCGCTTACCGTATC 597
QY 223 GlyPheLeu--SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProA 242
D 598 GGCTTCTTGTGCGCTTCATAGCCTTACTGGCTTGTATTGTCG----- 641
QY 242 spGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHisProA 262
D 642 -----CATGCGCGCGCGCTGTGTGTCGCGAG 666
QY 262 sphisProthr----- 265
D 667 GATGGCCGACGAGTCTCTGGGCCAAGACGCGCGCAGCAAGCGGCTCGTATGGCTGTG 726
QY 266 -----GlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuL 284
D 727 GTGGTGACAGCTGTCTTGTGCACTAGCTT---CCTGCTTTTCCATCATCAAGACAGCC 783
QY 284 euProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysG 304

```

```

Db 784 TACTTGGCTGTGCTCCAGCGCGGTGTCTTCCCTGTGCTGGAGACCTTCGCTGT 843
QY 304 LylLeuGlnAspMetGluAlaSerGlyGluCysGluGln----- 316
D 844 GCCTACAAAGGCACCTCGGCCCTTCGCCAGTGTCAACAGTGTCTCGGACCCCATCTCTTC 903
QY 317 -----LeuProGlnProSerProValLeuSerPheLysGlyGlyLysA 331
D 904 TACTTCACACACAGAGTTCGGCGGC-AAACCCACGATCTCTTACAGAGGCTCACAG- 961
QY 331 snArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProAlaGlyA 351
D 962 -----CCAAAGTGGA 971
QY 351 rGlysArg 353
D 972 GAGGCAGA 979

RESULT 4
AY400798 987 bp DNA linear GSS 15-DEC-2003
LOCUS Pan troglodytes P2RY6 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION Genomic survey sequence.
ACCESSION AY400798
VERSION AY400798.1 GI:39756787
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 987)
AUTHORS Clark,A.G., Glancowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 987)
AUTHORS Clark,A.G., Glancowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..987
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>987
/gene="P2RY6"
/locus_tag="HCM0677"

gene

ORIGIN
Alignment Scores:
Pred. No.: 8.2e-30 Length: 987
Score: 416.00 Matches: 115
Percent Similarity: 47.59% Conservative: 33
Best Local Similarity: 36.98% Mismatches: 128
Query Match: 21.49% Indels: 36
DB: 9 Gaps: 10

US-10-763-972-2 (1-360) x AY400798 (1-987)

QY 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGlyLys 48
D 10 GACAAATGGCAGCGCCAGGCTTGGGCTTCCACCCACCACTGTGTCTACAGCGAAG 69

```

QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheLeuGlyLeuProtein 68
 Db 70 TTCACGCAACTGCTGCTGCCACTGCTGTTATCGGGTCTGCTGGCTGGCTGGCTGGCTG 129
 QY 69 AsnGlyThrValLeuTyrHisSerTyrGlyGlnThrLysArgTyrSerCysAlaThrThr 88
 Db 130 AACATCTGTCTCATCCAGATCTCCAGTCCCGGGGGCTGACCGCGCGCGGTG 189
 QY 89 TyrLeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 107
 Db 190 TACACCTAACTTCTGCTGGCTGACCTGCTATATGCTGCTGCTGCTGCTGCTGCTC 249
 QY 108 IleThrTyrSerLeuAspArgTyrProPheGlyGluLeuLeuCysLysLeuValHis 127
 Db 250 TACAACCTATGCCCCAGGTATCACTGGCGCTTTGGGACTTGGCGCTGGCGCTGGCGC 309
 QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuTyrCysIleSerValHis 147
 Db 310 TTCTCTTTCTATGCGCAACCTGCGATGCGAGCATCTCTTCTCACTGCTGCTGCTCCAG 369
 QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166
 Db 370 CGCTACTGGGCATCTGCCACCGCTGGCGCTGGCGCTGGCGCAAAAGTGGGGCGCGCGCT 429
 QY 167 AlaTyrLeuGlyThrSerThrThrTyrAlaLeuValValLeuGlnLeuLeuProThrLeu 186
 Db 430 GCCTGGCTAGTGTGTAGCCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC 489
 QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGln 206
 Db 490 ATCTTGTCTGCCACGAGCATCCAGCGTAAACCGACTGCTGCTATGCTGCTGCTGCTG 549
 QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--- 225
 Db 550 GCCTGGCCACCCACTATATGCGCTACCGCATGGCTGCTCACTGCTCACTGCTGCTGCTG 609
 QY 226 -----SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAsp 242
 Db 610 CCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645
 QY 243 GlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAsp 262
 Db 646 ---GCGCGC-----CGCTGTGCGCGCAGGATGGCGCGCAGAG-----CCTGT- 686
 QY 263 HisProThrGlyValTyrProLeuHisPro-----TyrHisSerLeuLeuProHis 275
 Db 687 GCGCCAGGAGCGCGTGGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 746
 QY 276 ---AlaLeuPro-----TyrHisSerLeuLeuProHisHisLeuLeuSerAla 291
 Db 747 CATCAGCTTCTTACCTTTTACATCACCAAGACAGCGCTACCTGGCAGTGGCGCTCAACGCC 806
 QY 292 PheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSer 311
 Db 807 GGGCGTCCCTGCACTGATTGGAGGCTTTTGACGGCGCGCTTCAAAAGGCGCGCGCGGT 866
 QY 312 GlyGluCysGluGlnLeuProGlnProSerPro 322
 Db 867 TGCCAGTGGCAACAGCGTGGTGGACCCCATCCN 899

RESULT 5
 LOCUS CR624871
 DEFINITION full-length cDNA clone CS0D1035YK19 of Placenta Cot 25-normalized
 of Homo sapiens (human).
 ACCESSION CR624871
 VERSION CR624871.1 GI:50505678
 KEYWORDS HTC; CDS; cDNA
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1425)
 Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue
 2 (bases 1 to 1425)
 Genoscope.
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life technologies, a
 division of Invitrogen.
 Location/Qualifiers
 1. 1425
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1035YK19"
 /tissue type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

Alignment Scores:
 Pred. No.: 2,02e-29 Length: 1425
 Score: 414.50 Matches: 133
 Percent Similarity: 43.29% Conservative: 38
 Best Local Similarity: 33.67% Mismatches: 166
 Query Match: 21.41% Indels: 60
 DB: 3 Gaps: 10

US-10-763-972-2 (1-360) x CR624871 (1-1425)
 QY 8 SerArgGlySerArgSerGlySerArgArgGlyAlaLeuLeuGluGlyAlaSerArg 27
 Db 45 GCTGAGGCGAGAGAGTGGGAGCGCTCAGGTCTGGCTTCCCTTTGCCCCAACACACTGG 104
 QY 28 AspMetGluLysVal----- 32
 Db 105 GACCAACAGCGGTGTCAGGAGCAGAGGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTG 164
 QY 33 -----AspMetAsnThrSerGlnGluGlnGlyLeu----- 42
 Db 165 TGGGCGAGCGATGGAATGGGACAAATGGCAGCGGAGGCTCTGGGCTTGGCCACCCACC 224
 QY 43 CysGlnPheSerGlyLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePhe 62
 Db 225 TGTGTCTACCGCGAGAACTTCAAGCAACTGTGTGCTGCTGCCACCTGTGTATTTCGGCGGTG 284
 QY 63 IleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTyrGlnThrLysArg 82
 Db 285 GCGGTGGCGCTGCCCTGAAACATCTGTGTATATCCAGATCTGCACGTCGCCCGCGGCC 344
 QY 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuTyrValLeu 102
 Db 345 CTGACCGCGACGGCGGTGTACACCTAAACCTTGTCTGCTGCTGCTGCTGCTGCTGCTG 404
 QY 103 ---LeuProPheLeuIleIleThrTyrSerLeuAspArgTyrProPheGlyGluLeu 121
 Db 405 TCCCTGGCGCTGCTCATCTACAACTATGCCAAGGTGATCACTGGCCCTTTTGGGACTTC 464
 QY 122 LeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 141
 Db 465 GCTGCGCGCTGCTGCTGCTTCTCTTCTATGCAACCTGCGAGCGAGCATCTCTCTCTC 524
 QY 142 ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyr 161
 Db 525 ACCTGATCAGCTTCCAGCGCTACCTGGGCATCTGCCACCGCTGCTGCTGCTGCTGCTG 584

/note="unnamed protein product; PURINERGIC RECEPTOR P2Y,
G-PROTEIN COUPLED 2 (SPTR|Q9CPZ4, evidence: FASTY, 100%ID,
100%length, match=1119)
putative"
/codon start=1
/protein_id="BAB30719.1"
/db_xref="GI:12856589"
translation="MAADLEPNWSTINGWDELGYKCRNEDEKYLPLVSVGVVC
VLGLCLNVVALYELCLRLATWASITTYMPLHVASLSVAASPLILVYVYARGDHPFS
TVLCKLVRFLFTNLYCSILPLTICISVHRCLGLVRLPLSLWGRVARYARRVAAVVWL
VLACAPVLYFTVTSVGRFRTTCHDTSARELFSHFVSSVMGLLPLFVPSVILVCL
VIMARLLPAPVGTGGPRKRSVRTIALVFAUCFLPFLIRITRILYISFRSLDL
SCHTLNAINMAYKIRPLASANSCLDPLVFLAGRLVRFARFADKPEPTPTSPQARR
KLGLHPRNRTVKDLISVSSDDSRRTSTPAGSETKDIRL"

ORIGIN

Alignment Scores:
Pred. No.: 4,88e-29 Length: 2542
Score: 414.50 Matches: 132
Percent Similarity: 48.66% Conservative: 50
Best Local Similarity: 35.29% Mismatches: 155
Query Match: 21.41% Indels: 38
DB: 3 Gaps: 9

US-10-763-972-2 (1-360) x AK017378 (1-2542)

Qy 2 LeuSexIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeuLeu 21
Db 435 CTGAGCATCTC-----ACCACCTCAAGACGAGCGAGTGATCAG 473
Qy 22 LeuGluGly-AlaSerArgAspMetGluLysValAspMetAsnThrSer----- 37
Db 474 GTCCAGGGCAATGCGACGACCTGGACCCCTGGAAATAGCACCATCAATGGACCTGGGA 533
Qy 38 -----GlnGluGlnGlyLeu---CysGlnPheSerGluLysTyrLysGlnValTyrLeuSe 55
Db 534 GGGGACCAACTGGATACAGTGTCTTTCACAGGAGCTTCAAGTACGTGTGTGGC 593
Qy 55 rLeuAlaTyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHi 75
Db 594 CGTGTCTAATGGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 653
Qy 75 sSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyrValAsnLeuMetVa 95
Db 654 CTCTCTATGGCGCTCAAAACCTGGAACGCTCCACCATCATGTTCACCTGGCAGT 713
Qy 95 lAlaAspLeuLysTrpValLeu---LeuProPheLeuIleIleThrTyrSerLeuAspAs 114
Db 714 TTCGACTCTCTACGACGCTCCCTGCGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTG 773
Qy 114 PArgTrpPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLe 134
Db 774 CCACTGGCCATTTAGCAGCGTGTCTGCAAGCTGGTGTGGTGTGGTGTGGTGTGGTGTG 833
Qy 134 uTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHi 154
Db 834 CTACTGAGATCTCTCTCTCTACCTGATCAGCGTGCACCGGTGCTGGAGTCTCTGG 893
Qy 154 sProLeuCysSerLeuProTyrArgThrArgArgHisAlaTrpLeuGlyThrSerThr 174
Db 894 CCCTCTGACCTCCCTGGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 953
Qy 174 rTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyril 194
Db 954 GTGGGTGTGGTGTGGCTGCCAGGACCGCGTCTCTACTTGTCTACACCGACGCTGG 1013
Qy 194 eAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAl 214
Db 1014 GGGAAACCGGATCACTTGGCCATGACACTCGGCCCGGAGAGCTCTTTAGCCATTTTGG 1073
Qy 214 aTyrGlyIleValLeuThrLeuSerGlyPhe---LeuSerLeuLeuGlyHisPheGlyVa 233
Db 1074 TTACAGCTCCGTCATGTGGG-TCTGCTTTTGTGTGGTGTGGTGTGGTGTGGTGTGGT 1132

Qy 233 lLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGl 253
Db 1133 GTTACGTGCTTATGCCAGCGGGCTGCTCAACCGGCTTATGGACACAGGAGTCTGC 1192
Qy 253 nHisSerProSerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLe 273
Db 1193 CTCGGGCAACGGAAGTCTGTGGCCACCATTCGCTTGGTACTGGCGCTCTTCGCCCTCT 1252
Qy 273 uPheCysAlaLeuProTyrHisSer-----LeuLeuLeuProHisHisLeuLeuSe 290
Db 1253 GCTTCTGCTTTCACCTCAGCGGCACCTCTACTACTCTCTCTCG-----AT 1300
Qy 290 rAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAl 310
Db 1301 CACTTGACCTCAGCTGCCACACCTCAACGCGCATCAACATGGCATATAAGATCAACCCG 1360
Qy 310 aSerGlyGluCysGlnGlnLeuProGlnPro-SerProValLeuSerPheLysGlyL 330
Db 1361 CGTGTGGCAGCGCCCAACAGTTG-----TCTTGACCGGCTACTTCTTCTTCTGCGAGGC 1414
Qy 330 ysAsnArgValArgLeuLeuGlnLysLeuArg----- 340
Db 1415 AGAGACTTGTTCGCTTTCGCGAGATGCCAAGCCACCCAGCGAGCTACCCCGACCCAC 1474
Qy 341 -----GlnAsnLysLeuGlyGluHisProAlaGlyArg 351
Db 1475 AGGCTCGTGGCAAGTGGCGCTGCACAGGCGCTAACAAGA 1512
RESULT 7
AK005013 3001 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus adult male liver cDNA, RIKEN full-length enriched
DEFINITION library, clone:130001504 product:PURINERGIC RECEPTOR P2Y,
G-PROTEIN COUPLED 2, full insert sequence.
ACCESSION AK005013
VERSION AK005013.2 GI:26339698
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Makamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wachihi, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Nature 409, 685-690 (2001)
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
Nature 420, 563-573 (2002)
6 (bases 1 to 3001)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numata, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
On Dec 10, 2002 this sequence version replaced gi:12836638.
Please visit our web site (http://genome.gsc.riken.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGCGCGCACTCGAGTGTGTTTATTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGCGTCAAGTCAATTAATTAACCCGCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.
Location/Qualifiers
1..3001
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:1300015C04"
/db_xref="taxon:10090"
/clone="1300015C04"
/sex="male"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
477..1598
/note="unnamed protein product; PURINERGIC RECEPTOR P2Y, G-PROTEIN COUPLED 2 (SPTR)Q9CPZ4, evidence: FASTY, 100%ID, 100%length, match=1119)
putative"
/codon_start=1
/protein_id="BAB23746.1"
/db_xref="GI:12836639"
/translation="MAADLEPWNSTINGTWEGDELGYKCRFNEFKYLLPVSYGVVC VLGLNVALYIFLCRLTNASTYIMFELAVSDLSYASLPLVYYARVGDHWPFS TVLCKVRFYFTNGLIFLCISVHRCGLVRLHSLRWGRARYARVAVWVVL VLACQAPLYFTVTSVGRITCTDTSARELFSHFVYSSVLMGLLFAVPFYSVLVCL VIMARLLPAYGTGRTGPRKTSVRTILVAFALCFPLFHLFRTLYYSFSLDL SCTLNAINWAKYITPLASANSCLDPLVFLAGDLVSEFARDAKPEFTPTSPQARR KLGLRPNRVTRKDLSSVSDSRRTESTPAGSETKDIRL"
polyA_signal
2982..2987
/note="putative"
3001
polyA_site

ORIGIN
Alignment Scores:
Pred. No.: 6.29e-29 Length: 3001
Score: 414.50 Matches: 132
Percent Similarity: 48.68% Conservative: 50
Best Local Similarity: 35.29% Mismatches: 155
Query Match: 21.41% Indels: 38
DB: 3 Gaps: 9
US-10-763-972-2 (1-360) x AK005013 (1-3001)
QY 2 LeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeuLeu 21
DB 428 CTGAGCATCCTC-----ACCACCTCAAGAGCAGGAGCTGATCAG 466
QY 22 LeuGluGly-AlaSerArgAspMetGluLysValAspMetAsnThrSer----- 37
DB 467 GTCCAGGCAATGGCAGCAGACCTGGAACCTGGATAGCACCATCAATGGCAGCTGGGA 526
QY 38 -----GlnGlnGlnGlyLeu---CysGlnPheSerGluLysTyrLysGlnValTyrLeuSe 55
DB 527 GGGGACGAACTGGGATACAAAGTGTGCTTCAACGAGGACTTCAAGTACGTGCTGTGCC 586
QY 55 rLeuAlaTyrSerIleIlePheIleLeuGluGlyLeuProLeuAsnGlyThrValLeuTyrHi 75
DB 587 CGTGTCCTATGGCGTGTGCTCGGTGTGCTGCTGAACTGCTGGGCTCTCTATAT 646
QY 75 sSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetVa 95
DB 647 CTTCTATGCGCCTCAAAACCTGGAACCTGCAACCTCACACCTACATGTTTACCTGGCAGT 706
QY 95 lAlaAspLeuLeuTyrValLeu---LeuProPheLeuIleIleThrTyrSerLeuAspAs 114
DB 707 TTGGGACTCTCTTACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 766
QY 114 pArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLe 134
DB 767 CCACTGGCCATTAGCACGGTGCTCTGCAAGCTGGTGGTGTCTCTTCTTACACCAACCT 826
QY 134 uTyrGlySerIleLeuLeuLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHi 154
DB 827 CTATCTGACGATCTCTTCTCTCCTCCTGATCAGCGTGCACCGCTGCTGGGAGTCTCTGC 886
QY 154 sProLeuCysSerLeuProTyrArgThrArgHisAlaTrpLeuGlyThrThrTh 174
DB 887 CCTCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 946
QY 174 rTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrI 194
DB 947 GTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1006
QY 194 eAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGlnAsnPheAspArgLeuPheAl 214
DB 1007 GGGAAACCGGATCACTTGCCATGACACCTGCGCCGAGAGCTCTTTAGCCATTTTGTGGC 1066
QY 214 aTyrGlyIleValLeuThrLeuSerGlyPhe---LeuSerLeuLeuGlyHisPheGlyVa 233
DB 1067 TTACAGCTCCGTCATGCTGGG-TCTGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1125
QY 233 lLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluProHisgluAspArgGl 253
DB 1126 GTTACGTCTTATGGCAGCGCGGTGCTCAAAACCGGCTTATGGGACCAAGAGGCTCTGC 1185
QY 253 nHisSerProSerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLe 273
DB 1186 CTCGGGCAACGCGAGTCTGTGCGCACCATTTGCTTGGTACTGGCCGCTCTTCCCTCTCT 1245
QY 273 uPheCysAlaLeuProTyrHisSer-----LeuLeuLeuProHisHisLeuLeuSe 290
DB 1246 GCTTCTGCTTCTCCACCTCAACGCGCACCTCTACTACTCTCTCTCTCTCTCTCTCTCTCT 1293

/note="putative"

```

QY 290 xAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAl 310
Db 1294 CACTTGACCTCAGCTGCCACACCTCAACGCCATCAACATGCGCATATAGATCAACCGGC 1353
QY 310 aSerGlyGluCysGluGlnLeuProGlnPro-SerProValLeuSerPheLeysGlyGlyL 330
Db 1354 CGTGCGCCAGCCCAACAGCTTG-----TCCTTGACCCGGTACTCTACTCTCTGGCAGGC 1407
QY 330 yAAsnArgValArgLeuLeuGlnLysLeuArg----- 340
Db 1408 AGAGACTTGTCGCTTGCCGCTTGCCGAGATGCCAAGCCACCCAGGAGCTACCCCGCCAC 1467
QY 341 -----GlnAsnLysLeuGlyGluHisProAlaGlyArg 351
Db 1468 AGGCTCGTCGCAAGCTGGGCGCTGCACAGCGCTAAACAGA 1505

RESULT 8
AY400797 987 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens P2RY6 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY400797
VERSION AY400797.1 GI:39756786
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 987)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene tricos
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 987)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
location 1..987
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..5987
/gene="p2ry6"
/locus_tag="HCM0677"

gene

ORIGIN
Alignment Scores:
Pred. No.: 2,53e-29 Length: 987
Score: 411.00 Matches: 125
Percent Similarity: 48.30% Conservative: 31
Best Local Similarity: 38.70% Mismatches: 147
Query Match: 21.23% Indels: 22
DB: 9 Gaps: 8

US-10-763-972-2 (1-360) x AY400797 (1-987)

QY 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48
Db 10 GACAATGCACAGCCAGGCTCTGGCTTGCCACCCACCACTGTGTCTACCGCAGAAC 69
QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 68

```

```

Db 70 TTCAGCAACTGCTGCTGCCACCTGTGTATTCCGGCGGTGCTGGCGGCTGCCCTGCCGCTG 129
QY 69 AsnGlyThrValLeuTyrHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88
Db 130 AACATCTGTGTCAATCCAGATCTGCACGCTCCGCGCGGCGCTGACCCGACGGCGCTG 189
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuLe 107
Db 190 TACACCTTAACCTTGTCTGGCTGACCTGTATATGCTGCTGCTGCTGCTGCTGCTGCTCATC 249
QY 108 IleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127
Db 250 TACAACTATGCCCAAGGTGATCACTGGCCCTTTGGCGACTTGGCTGCCCTGGTCCGCG 309
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 310 TTCCTCTTCTATGCAACCTGCACGCGAGCATCTCTTCTCTACCTGCATCAGCTTCCAG 369
QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166
Db 370 CGCTACCTTGGCATCTGCCACCCCTGGCCCCCTGGCACAAACGTGGGGGCGCGCGGCT 429
QY 167 AlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu 186
Db 430 GCCTGGCTAGTGTGTAGCGCTGTGGCTGGCGCTGACAAACCCAGTAGCTGCCTGCCACAGC 489
QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTrpAspMetThrSerGln 206
Db 490 ATCTTCGCTGCCACAGGATCCACGCTGTAACCGACTGTCTGTATGACTCAGCCCGGCT 549
QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyLeValLeuThrLeuSerGlyPheLeu--S 226
Db 550 GCCCTGGCCACCCACTATATGCCCTATGGCATGGCTCTCACTGTCTGCTGCTGCTGCTG 609
QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246
Db 610 CCCTTTGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
QY 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
Db 667 GATGCCCC-----GGCAGACCTTGTGGCCACAGAGCGGCTGCAAGCGCGCGCATG 720
QY 266 lyValTrp-----ProLeuHisProPhePheCysAlaLeuProTyrHisSer 281
Db 721 GCCG-TGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 776
QY 282 LeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
Db 777 GACAGCCTACCTGCGCAGTGGCTGACCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 836
QY 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSer 321
Db 837 TGCAGCGCGCTACAAAGCAGCGCGCGCTTTGGCAGTGCACAGCGCTGCTGCTGCTGCTG 896
QY 322 ProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuGlnLys 338
Db 897 CCTCTTCTAC--TTCACCCAGAGAAGTTCGCCCGCGGACCCACATGAGCTCCTACAGAA 954

339 Leu 339
955 CTC 957

RESULT 9
CR612681 1515 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DI044YA15 of Placentia Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR612681
VERSION CR612681.1 GI:50493488
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1515)
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1515)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1044YA15"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
source

Alignment Scores:
Pred. No.: 4.87e-29 Length: 1515
Score: 411.00 Matches: 126
Percent Similarity: 45.85% Conservative: 34
Best local Similarity: 36.10% Mismatches: 152
Query Match: 21.23% Indels: 39
DB: 3 Gaps: 9

US-10-763-972-2 (1-360) x CR612681 (1-1515)

QY 33 AspVetAnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48
DB 271 GACAATGGCAGGCGGCTCTGGGCTTCCACCCACCACTGTCTACCGCGAGAAC 330
QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIlePheLeuGluLeuProLeu 68
DB 331 TTCAGGCACTGCTGCTGCCACCTGTATTCGGCGTGTGGCGCTGGCGCTGCGCTG 390
QY 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThr 88
DB 391 AACATCTGTCTATTACCCAGATCTGCACCTCCCGCGCGGCTGTACCGCGCGCTG 450
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu--LeuProPheLeu 107
DB 451 TACACCTAACCTTGTCTGGCTGACCTCTATATGCTGCTCCCTGCTGCTCATC 510
QY 108 IleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127
DB 511 TACAATATGCCAAGGTATCACTGGCCCTTTGGCGACTTGGCGCTGCTGCTGCTG 570
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
DB 571 TTCTCTCTTCATGCCAACCTGCGCAGCGAGCATCTTCTTCACTGATCATCTCCAG 630
QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg--ThrArgArgHis 166
DB 631 CGTACTCTGGCATCTGCCACCGCTGGCCCTGGCCCAAGCTGGGGCGCGCGGCT 690
QY 167 AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeuLeuProThrLeu 186
DB 691 GCTGTGCTAGTGTGTAGCGGTGTGGCTGACCAACCGAGTGGCTGCGCCACAGCC 750
QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206

DB 751 ATCTTCTGCTGCCACAGGCATCCAGCGTAACCGCACTGTCTGTATGACCTCAGCCGCT 810
QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226
DB 811 GCCCTGGCCACCACTATATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246
DB 871 CCCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927
QY 246 lYGluproHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
DB 928 GATGCCCC-----GGCAGAGCCTGTGCCCGAGCGCGCTGGCAAGCGGCCCGCATG 981
QY 266 lYValTrp-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281
DB 982 GCGC-TGGTGGTGGCTGCTGCTTGGCATCAGCTT---CCTGGCTTTTCATCATCACC 1037
QY 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
DB 1038 GACAGCTTACCTGGCAGTGCCTGCACGCCGGCGGTCCCTGCTGCTGCTGCTGCTG 1097
QY 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSer 321
DB 1098 TGCAGCGGCTTACAAAGCAGCGCGCTTGGCGAGTGCAGTGCAGCGCTGCTGCGAC 1157
QY 322 ProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArgGln 341
DB 1158 CCT-----CTTCTACT-TCACCCAGAAGA 1180
QY 342 AsnLysLeuGlyGluHis-----ProAlaGlyArgLysArg 353
DB 1181 AGTTCGCGCGCGACCATCATGAGCTCTACAGAACTCAGACCAATGGCAGGCGAG 1240
QY 354 CysProGlyLeuAsnArgSer 360
DB 1241 GTCGCTGAGTCTCCAGGTCC 1261

RESULT 10

CR618945 1532 bp mRNA linear HTC 21-JUL-2004
LOCUS
DEFINITION
full-length cDNA clone CSOD1042YA16 of Placenta Cot 25-normalized
of Homo sapiens (human).

ACCESSION
CR618945

VERSION
CR618945.1 GI:50499752

KEYWORDS
HTC; CNSL7; cDNA.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE
Full-length cDNA libraries and normalization

JOURNAL
Unpublished

REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1532)
Genoscope.
Direct Submission

AUTHORS
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :

JOURNAL
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1..1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source

/clone="CSODI042YA16"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="PCMVSPORT_6"

ORIGIN

Alignment Scores:
 Pred. No.: 4,96e-29 Length: 1532
 Score: 411.00 Matches: 126
 Percent Similarity: 45.85% Conservative: 34
 Best Local Similarity: 36.10% Mismatches: 152
 Query Match: 21.23% Indels: 39
 DB: 3 Gaps: 9

US-10-763-972-2 (1-360) x CR618945 (1-1532)

QY 33 AspMetAsnThrSerGlnGluGlyLeu-----CysGlnPheSerGluLys 48
 Db 271 GAAATGGCAGCCAGGCTCTGGGCTTCCACCCACACCTGTGTCTACCCGAGAAC 330
 QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeu 68
 Db 331 TTCAAGCAACTGCTGCTGCCACCTGTATTCCGGCGTGTGGCGCTGCGCTGCGCTG 390
 QY 69 AsnGlyThrValLeuThrPheHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88
 Db 391 AACATCTGTCTCATACCCAGATCTGCACGTCCCGCGCGCCCTGACCCGACGCGCGTG 450
 QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 107
 Db 451 TACACCTAAACCTTGTCTGGCTGACCTGTATATCCCTGCTCTGCTGCTGCTCATC 510
 QY 108 IleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127
 Db 511 TACAACATATGCCAAGGTGATCACTGGCCCTTTGGGACCTTCGCTGCGCTGCTGCGC 570
 QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
 Db 571 TTCTCTTCTATGCCAACCTTGCCTGGCTGACCTGTATATCCCTGCTCTGCTGCTCATC 630
 QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166
 Db 631 CGCTACCTGGGCACTGCCACCGCTGCCCGCTGGCCACAAAGTGGGGCGCGCGGCT 690
 QY 167 AlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu 186
 Db 691 GCCTGGCTAGTGTGTAGCGGTGTGGTGGCTGACACCCAGCTGCTGCGCCACAGCC 750
 QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206
 Db 751 ATCTTCGCTGCCAGGCACTCCAGCGTAACCGCACTGTCTGTATGACCTCAGCGCGCT 810
 QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226
 Db 811 GCCCTGGCCACCCACTATATCCCTATGGCATGCTCTCACTGTCTGCTGCTGCTGCTG 870
 QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246
 Db 871 CCCTTGTGCGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927
 QY 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
 Db 928 GATGCCCC-----GGCAGAGCTGTGGCCGAGGCGGTGGCAAGGGCGGCGCGCATG 981
 QY 266 lyValTrp-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281
 Db 982 GCGG-TGGTGTGGCTGTGCTTGTGCATCAGCTT---CTGCTTTTTCATCATCACCAA 1037
 QY 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
 Db 1038 GACACCTACTGCGAGTGGCTCGACCGCGGGCGTCCCTGCTGCTGCTGCTGCTGCTGCT 1097
 QY 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSer 321

Db 1098 TCCAGCGGCTTACAAAGCAGCAGCGCGCTTTTCCAGTGCCAAAGCGTGTGGAGCCCAT 1157
 QY 322 ProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArgGln 341
 Db 1158 CCT-----CTTCTACT--TCACCCAGAGA 1180
 QY 342 AsnLysLeuGlyGluHis-----ProAlaGlyArgLysArg 353
 Db 1181 AGTTCCGCGCGCAGCCACATGAGTCTTACAGAAACTCACAGCAATGGCAGAGGCAGG 1240
 QY 354 CysProGlyLeuAsnArgSer 360
 Db 1241 GTCGCTGAGTCTCTCCAGTCC 1261

RESULT 11
 LOCUS CR626266 1564 bp mRNA linear HTC 21-JUL-2004
 DEFINITION full-length cDNA clone CSODI038YC15 of Placenta Cot 25-normalized of Homo sapiens (human).
 ACCESSION CR626266
 VERSION CR626266.1 GI:50507073
 KEYWORDS HTC; CNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1564)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
 REFERENCE 2 (bases 1 to 1564)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequenage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 COMMENT - Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
 FEATURES
 source Location/Qualifiers
 1..1564 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI038YC15"
 /tissue_type="Placenta Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:
 Pred. No.: 5.12e-29 Length: 1564
 Score: 411.00 Matches: 126
 Percent Similarity: 45.85% Conservative: 34
 Best Local Similarity: 36.10% Mismatches: 152
 Query Match: 21.23% Indels: 39
 DB: 3 Gaps: 9

US-10-763-972-2 (1-360) x CR626266 (1-1564)

QY 33 AspMetAsnThrSerGlnGluGlyLeu-----CysGlnPheSerGluLys 48
 Db 302 GAAATGGCAGCCAGGCTCTGGGCTTCCACCCACACCTGTGTCTACCCGAGAAC 361
 QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeu 68
 Db 362 TTCAAGCAACTGCTGCTGCCACCTGTATTCCGGCGTGTGGCGCTGCGCTGCGCTG 421
 QY 69 AsnGlyThrValLeuThrPheHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88


```

Db 422 AACATCTGTGATACCCAGATCTGCAGTCCCGCGGGCCCTGACGCCGACGCGCGCTG 481
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 107
Db 482 TACACCTTAACCTTGCTGTGCTGACCTGTATATGCTCTCTCCCTGCGCTGCTATC 541
QY 108 lleThrTyrSerLeuAspArgTTPProPheGlyGluLeuLeuCysLysLeuValHis 127
Db 542 TACAACCTATGCCAAGGTGATCACTGGCCCTTTGGCGACTTCGCGCTGCGCTGGTCCG 601
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 602 TTCCTCTTCTATGCGAACTGCAGCGAGCATCTCTTCTCACTGCATCAGTTCACAG 661
QY 148 GlnPheLeuGlyValCysHisProLeuLeuCysSerLeuProTyrArg---ThrArgArgHis 166
Db 662 CGCTACCTGGCATCTGCCACCCGCTGGCCCTGGCCACAACTGGGGCGCGCGGCT 721
QY 167 AlaTTPLeuGlyThrSerThrTTPAlaLeuValValLeuGlnLeuLeuProThrLeu 186
Db 722 GCCTGGCTAGTGTGTAGCGGTGTGGCTGGCGGTGACAAACCCAGTGCCTGCCACAGCC 781
QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTTPTrpAspMetThrSerGln 206
Db 782 ATCTTGCTGTCACAGGATCCAGGTAACCGCACTGTCTGTATGACCTCAGCCGCGCT 841
QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226
Db 842 GCCTGGCCACCCACTATATCCCTATGGCATGGCTCTCACTGTGCTGCTGCTGCTG 901
QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246
Db 902 CCCCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 958
QY 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
Db 959 GATGGCCC-----GGCAGAGCCCTGTGGCCCGAGGCGGTGGCAAGGCGCGCGCATG 1012
QY 266 lyValTTP-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281
Db 1013 GCCG-TGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
QY 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
Db 1069 CACAGCCTACCTGGCAGTGGCTCGACGCGGGGGTCCCTGCACTGTATTGGAGGCTT 1128
QY 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSer 321
Db 1129 TGCAGCGCGCTTACAAAGGCACGCGCGCTTTGGCAGTGCCCAACAGCGGTGTGGACCCCA 1188
QY 322 ProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArgGln 341
Db 1189 CCT-----CTTCTACT-TCACCCAGAGA 1211
QY 342 AsnLysLeuGlyGluHis-----ProAlaGlyArgLysArg 353
Db 1212 AGTTCCCGCGCGACCATGAGCTCTCAGAAACTCAGAGCAAAATGGCAGGCGAGG 1271
QY 354 CysProGlyLeuAsnArgSer 360
Db 1272 GTCGCTGAGTCTCCAGGTCC 1292

```

RESULT 12

```

CR626754
LOCUS full-length cDNA clone CSOD1024YE14 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR626754
VERSION CR626754.1 GI:50507616
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1589)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 1589)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
EP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1024YE14"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 5,246-29 Length: 1589
Score: 411.00 Matches: 126
Percent Similarity: 45.85% Conservative: 34
Best Local Similarity: 36.10% Mismatches: 152
Query Match: 21.23% Indels: 39
DB: 3 Gaps: 9
US-10-763-972-2 (1-360) x CR626754 (1-1589)

```

```

QY 33 AspMetAsnThrSerGlnGlnGlyLeu-----CysGlnPheSerGluLys 48
Db 341 GACAATGCCACAGCGCGCTGTGGCTTGCCACCCACCTGTGCTACCGCGAGAAC 400
QY 49 TyrLysGlnValTyrSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeu 68
Db 401 TTCAGCACTGCTGCTGCCACCTGTGTATTGGCGGTGCTGGCGCTGCGCTG 460
QY 69 AsnGlyThrValLeuTTPHisSerTTPGlyGlnThrLysArgTTPSerCysAlaThrThr 88
Db 461 AACATCTGTGATATACCCAGATCTGCCAGTCCCGCGCGCTGACCCGACGCGCGCTG 520
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 107
Db 521 TACACCTTAACCTTGCTGTGCTGACCTGTATATGCTCTCTCCCTGCGCTGCTATC 580
QY 108 lleThrTyrSerLeuAspArgTTPProPheGlyGluLeuLeuCysLysLeuValHis 127
Db 581 TACAACCTATGCCAAGGTGATCACTGGCCCTTTGGCGACTTCGCGCTGCGCTGGTCCG 640
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 641 TTCCTCTTCTATGCGAACTGCAGCGAGCATCTCTTCTCACTGCATCAGTTCACAG 700
QY 148 GlnPheLeuGlyValCysHisProLeuLeuCysSerLeuProTyrArg---ThrArgArgHis 166
Db 701 CGCTACCTGGCATCTGCCACCCGCTGGCCCTGGCCACAACTGGGGCGCGCGGCT 760
QY 167 AlaTTPLeuGlyThrSerThrTTPAlaLeuValValLeuGlnLeuLeuProThrLeu 186
Db 761 GCCTGGCTAGTGTGTAGCGGTGTGGCTGGCGGTGACAAACCCAGTGCCTGCCACAGCC 820
QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTTPTrpAspMetThrSerGln 206

```



```

Db 821 ATCTTCGCTGCCACAGGATCCAGGCTAACCAGCACTGTCTGTATGACCTACGCCGCGCT 880
Qy 207 GluAsnPhaAspArgLeuPheAlaTyGlyIleValLeuThrLeuSerGlyPheLeu--S 226
Db 881 GCCTGGCCACCACTATATGCCATATGCCATATGCCATATGCCATATGCCATATGCCAT 940
Qy 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnLeuProAlaAlaArgG 246
Db 941 CCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
Qy 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
Db 998 GATGGGCC-----GGCAGAGCCTGTGGCCAGGCGGCTGGCAGGCGGCGCCGCGATG 1051
Qy 266 lyValTrp-----ProLeu-HisProLeuPheCysAlaLeuProTyHisSer 281
Db 1052 GCCG-TGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1107
Qy 282 LeuLeuLeuProHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
Db 1108 GAGAGCCTACTGGCGTGGCTGCGAGCGCGGCGTCCCTGACACTGATGATGGAGCCCTT 1167
Qy 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGlyCysGlnLeuProGlnProSer 321
Db 1168 TGCAGGCGCTACAAAGGACGCGGCGCTTGGCAGTGGCCACAGCGTGTGGACCCCAT 1227
Qy 322 ProValLeuSerPheGlyGlyGlyLysAsnArgValArgLeuGlnLysLeuArgGln 341
Db 1228 CCT-----CTTCTACT-TCACCCAGAGA 1250
Qy 342 AsnLysLeuGlyGluHis-----ProAlaGlyArgLysArg 353
Db 1251 AGTTCGCGCGGACACACATGAGCTCTACAGAACTCAGAGCAATGCGAGGACAG 1310
Qy 354 CysProGlyLeuAsnArgSer 360
Db 1311 GTGGCTGAGTCTCCAGGTCC 1331

RESULT 13
LOCUS CR605588 1592 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI040YB02 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR605588
VERSION CR605588.1 GI:50486395
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1592)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope 2 (bases 1 to 1592)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1592
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="CS0DI040YB02"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores: 5.26e-29 Length: 1592
Pred. No.: 411.00 Matches: 126
Score: 45.85% Conservative: 34
Best Similarity: 36.10% Mismatches: 152
Best Local Similarity: 21.23% Indels: 39
Query Match: 3 Gaps: 9
DB: 9

US-10-763-972-2 (1-360) x CR605588 (1-1592)
Qy 33 AspMetAsnThrSerGlnGlnGlnGlyLeu-----CysGlnPheSerGlyLys 48
Db 349 GACAAATGGCAGCAGCCAGGCTCTGGGCTTGCCACCCACCTGTCTACCGCGAGAAC 408
Qy 49 TyrlsGlnValTyrlsLeuSerLeuAlaTyrlsSerIlePheIleLeuGlyLeuProLeu 68
Db 409 TTCAGCAACTGCTGCTGCCACCTGTGTATTCGGCGGTGGCGGCTGCGCTCCGCGTG 468
Qy 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThr 88
Db 469 AACATCTGTGTCTATTACCCAGATCTGCACGTCCGCGGCGCCCTGACCCGACGCGCGTG 528
Qy 89 TyrlsValAsnLeuMetValAlaAspLeuLeuTyrlsValLeu---LeuProPheLeu 107
Db 529 TACACCTTAACCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 588
Qy 108 IleThrTrpSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127
Db 589 TACAACTATGCCAAGGTGATCACTGGGCTTGGCGACTTGGCGCTGCGCTGCGCTGCGC 648
Qy 128 PheLeuPheTyrlsAsnLeuTyrlsSerIleLeuLeuLeuThrCysIleSerValHis 147
Db 649 TTCCTCTTCTATGCAACCTGCACGCGAGCATCTCTTCTCCTACCTGATCAGCTTCAG 708
Qy 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrlsArg---ThrArgHis 166
Db 709 CGCTACTTGGGATCTGCCACCGCTGCGCCCTGGCACAACAGTGGGCGCGCGGCT 768
Qy 167 AlaTrpLeuGlyThrSerThrTrpAlaLeuValLeuGlnLeuLeuProThrLeu 186
Db 769 GCCTGGCTAGTGTGTAGCGGTGTGGCGGTGACCAACCCAGTGGCTGCGCCAGCGC 828
Qy 187 AlaPheSerHisThrAspTyrlsAsnGlyGlnMetIleTrpTyrlsAspMetThrSerGln 206
Db 829 ATCTTGGCTGCCACAGGATCCAGCGTAAACCGCACTGTCTGCTATGACTCAGCCGCGCT 888
Qy 207 GluAsnPhaAspArgLeuPheAlaTyGlyIleValLeuThrLeuSerGlyPheLeu--S 226
Db 889 GCCTGGCCACCACTATATGCCATATGCCATATGCCATATGCCATATGCCATATGCCAT 948
Qy 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnLeuProAlaAlaArgG 246
Db 949 CCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
Qy 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
Db 1006 GATGGGCC-----GGCAGAGCCTGTGGCCAGGCGGCTGGCAGGCGCGCGCGCATG 1059
Qy 266 lyValTrp-----ProLeu-HisProLeuPheCysAlaLeuProTyHisSer 281
Db 1060 GCCG-TGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
Qy 282 LeuLeuLeuProHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
Db 1116 GACAGCCTTACCTGGCAGTGGCTCGAGCGCGGCGCTCCCTGCACTGTATTGGAGCCCTT 1175
Qy 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnLeuProGlnProSer 321

```


ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1638)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
2 (bases 1 to 1638)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1..1638
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS001040YP17"
/tissue_type="placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 5,49e-29 Length: 1638
Score: 411.00 Matches: 126
Percent Similarity: 45.85% Conservative: 34
Best Local Similarity: 36.10% Mismatches: 152
Query Match: 21.23% Indels: 39
DB: 3 Gaps: 9
US-10-763-972-2 (1-360) x CR617080 (1-1638)
QY 33 AsnMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGlnLys 48
DB 374 GACAATGGCAGCAGCCAGGCTCTGGGCTTCCACCCACACCTGTGTCTACCCGAGAAC 433
QY 49 TyrlsGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeu 69
DB 434 TTCAAGCACTGCTGCTGCCACCTGTATTGGGGGTGCTGGCGCTGGCCCTGGCGCTG 493
QY 69 AsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88
DB 494 AACATCTGTGTCAATTACCCAGATCTGCACGTCCCGCGGCCCTGACCCGCGCGCTG 553
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 107
DB 554 TACACCTAAACCTTGTCTGGTGACCTGCTATATGCTTCCCTGCTCCCTGCTCATC 613
QY 108 IleThrTyrSerLeuAspArgTrpProPheGlyGlnLeuLeuCysLysLeuValHis 127
DB 614 TACAACATATGCCAAGGTGATCACTGGCCCTTTGGCGACTTCGCGCTGGCCCTGGTCCGC 673
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuThrCysIleSerValHis 147
DB 674 TTCCTCTTCTATGCAACCTGCGGAGGAGATCTCTTCCTACCTGATCAGCTTCCAG 733
QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgHis 166
DB 734 CGCTACCTGGGCATCTGCCACCCGCTGGCCCTGGCACAAACGTGGGGCGCGCGGCT 793
QY 167 AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu 186
DB 794 GCCTGGCTAGTGTGTAGCCGTGTGGCTGGCGGTGACAAACCCAGTCCCTGCCACAGCC 853

QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206
DB 854 ATCTTCGCTGCCACAGCATCCAGCGTAACCGCACTGTCTGTCTATGACCTCAGCCGCGCT 913
QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226
DB 914 GCCCTGGCCACCACTATATGCCCTATGGCATGGCTCTCACTGTATCGCGTTCCTGCTG 973
QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgG 246
DB 974 CCCTTGCTGCCCTGTGGCTGTACTGTCTCTGGCGCTG---CCGCTGTGGCGCAG 1030
QY 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266
DB 1031 GATGCGCC-----GGCAGAGCCTGTGGCCAGGAGCGGTGGCAAGCGCGCGCATG 1084
QY 266 lyValTrp-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281
DB 1085 GCGG-TGTGTGTGGCTGTGGCTTTGGCATCAGTT---CCTGCTTTTACATCACCAA 1140
QY 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301
DB 1141 GACAGCCTTACCTGGCAGTGGCTCGACGCGGGGCTCCCTGCACTGTATTGGAGGCTT 1200
QY 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnGlnLeuProGlnProSer 321
DB 1201 TCCAGCGGCTTACAAAGGCACGCGCGCTTTGCCAGTGCCCAACAGCGTGTGGACCCCAT 1260
QY 322 ProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArgGln 341
DB 1261 CCT-----CTTCTACT-TCACCCAGAAGA 1283
QY 342 AsnLysLeuGlyGluHis-----ProAlaGlyArgLysArg 353
DB 1284 AGTTCCGCGCGGCGACCATGAGCTCTTACAGAAACTCACAGCCAAATGGCAGAGGCGG 1343
QY 354 CysProGlyLeuAsnArgSer 360
DB 1344 GTCGCTGAGTCTCTCCAGGTCC 1364

Search completed: November 10, 2004, 13:16:13
Job time : 3872 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 04:41:06 ; Search time 603 Seconds
(without alignments)
9419.359 Million cell updates/sec

Title: US-10-763-972-1

Perfect score: 1082

Sequence: 1 atgcgtccattttgtctcc.....ccagggttgacagatctgg 1082

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001s:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1082	100.0	1082	6	Abq76000 Human GPC
2	1082	100.0	1156	10	Adb99257 Human p2Y
3	1082	100.0	1334	10	Adc86166 Human GPC
4	1082	100.0	3143	6	Aas17746 Human gen
5	1059	97.9	1108	6	Adi16629 Human NOV
6	1059	97.9	1108	12	Adn42283 Human CDN
7	1051	97.1	1063	6	Aas98135 Human DNA
8	1006	93.0	1017	10	Adc25999 Human pur
9	1006	93.0	1017	10	Adc25999 Human pur
10	1002.8	92.7	1020	6	Adb79300 Human GPC
11	1002.8	92.7	1076	6	Adb29667 Human G-p
12	1002.8	92.7	1076	12	Ado30395 Human GPC
13	996.4	92.1	1002	4	Aah73516 Human G p
14	850	78.6	850	6	Aas17747 Human P2Y
15	606.2	56.0	906	12	Ado30397 Mouse GPC
16	510	47.1	510	4	Aah50975 Human nGP
17	510	47.1	510	6	Abs70208 DNA encod
18	396.4	36.6	681	4	Aas30774 Human CDN
19	370	34.2	585	4	Aas30783 Human CDN
20	235.8	21.8	1163	12	Adp49190 Chick P2Y
21	227.8	21.1	1134	12	Ado30017 Human GPC

22	227.8	21.1	2025	4	AAD04980	Aad04980 Human pur
23	227.8	21.1	2025	6	ABZ35611	Abz35611 Human gen
24	227.8	21.1	2025	8	ABZ42713	Abz42713 Human pur
25	227.8	21.1	2025	10	ACA56884	Aca56884 Human sig
26	227.8	21.1	2025	12	ADI56680	Adi56680 Human pol
27	227.8	21.1	2025	12	ADP49188	Adp49188 Human P2Y
28	226.2	20.9	2118	6	ABZ35320	Abz35320 Human gen
29	224.6	20.8	2547	12	ADP21380	Adp21380 Gene P2RY
30	223.6	20.7	1163	4	AAD04981	Aad04981 Human pur
31	222	20.5	551	4	AAS42837	Aas42837 Human G p
32	202.2	18.7	1122	12	ADO30307	Ado30307 Mouse GPC
33	202.2	18.7	1842	2	AAQ88134	Aaq88134 Human P20
34	202.2	18.7	2138	6	AAD32944	Aad32944 Mus muscu
35	202.2	18.7	2138	6	AAD32937	Aad32937 Mus muscu
36	187.8	17.4	2011	10	ADB52840	Adb52840 Primary r
37	178.6	16.5	1098	8	ABZ42771	Abz42771 Human pyr
38	178.6	16.5	1098	12	ADO30018	Ado30018 Human GPC
39	178.6	16.5	1428	12	ADP21381	Adp21381 Gene P2RY
40	178.6	16.5	1429	2	AAT74321	Aat74321 Human P2Y
41	178.6	16.5	1429	10	ACA56470	Aca56470 Human sig
42	178.6	16.5	1429	12	ADI56266	Adi56266 Human pol
43	178.6	16.5	1429	12	ADP49192	Adp49192 Human P2Y
44	173.8	16.1	1651	10	ADE40462	Ade40462 Human pyr
45	170	15.7	234	4	AAS30782	Aas30782 Human CDN

ALIGNMENTS

RESULT 1
ABQ76000
ID ABQ76000 standard; DNA; 1082 BP.
XX
AC ABQ76000;
DT 17-OCT-2002 (first entry)
XX
DE Human GPCR designated PFI-020 encoding sequence.
XX
KW Human; GPCR; G-protein coupled receptor; antidepressant; neuroleptic;
KW gene therapy; therapeutic; mood; depression; arousal; eating; sleeping;
KW disorder; PFI-020; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1080
FT /tag= a
FT /partial
FT /product= "GPCR designated PFI-020"
FT /note= "no stop codon given"

EP1215214-A1.
19-JUN-2002.
04-DEC-2001; 2001EP-00310137.
18-DEC-2000; 2000GB-00030855.
17-JAN-2001; 2001GB-00001222.
(PFI2) PFIZER LTD.
(PFI2) PFIZER INC.
Fidock MD;
WPI, 2002-510798/55.
P-PSDB; ABB98145.
New polynucleotide encoding G protein-coupled receptor PFI-020, useful
e.g. for treating eating and sleeping disorders and for identifying
specific modulators.

Claim 1 (c); Page 11; 23pp; English.

PS The invention relates to an isolated polynucleotide encoding a novel
 CC polypeptide belonging to the class of proteins known as G-protein coupled
 CC receptors (GPCRs). The activity of proteins of the invention may be
 CC described as, antidepressant and neuroleptic. Polynucleotides of the
 CC invention are used for recombinant expression of the G protein-coupled
 CC receptor (PFI-020) polypeptides, to create transgenic animals, as source
 CC of primers, probes, antisense sequences and ribozymes and in gene
 CC therapy. Therapeutic agents of the invention can be used to treat a wide
 CC range of disorders, particularly mood disorders, depression or arousal,
 CC especially eating and sleeping disorders. The current sequence represents
 CC a coding sequence for a human GPCR designated PFI-020

XX Sequence 1082 BP; 218 A; 318 C; 277 G; 269 T; 0 U; 0 Other;

Query Match 100.0%; Score 1082; DB 6; Length 1082;
 Best Local Similarity 100.0%; Pred. No. 3.9e-285;
 Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTGTCATTTGCTTCCTTCAGGGGAGCAGAGAGCGGAGCGCTGCTGAGCTCTG 60
 Db 1 ATGCTGTCATTTGCTTCCTTCAGGGGAGCAGAGAGCGGAGCGCTGCTGAGCTCTG 60
 Qy 61 CTCCTGAGGAGGCTCCCGGACATCGAGAGGTGGACATGAATACATCACAGGAACAA 120
 Db 61 CTCCTGAGGAGGCTCCCGGACATCGAGAGGTGGACATGAATACATCACAGGAACAA 120
 Qy 121 GGTCTCTGCCAGTTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
 Db 121 GGTCTCTGCCAGTTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
 Qy 181 ATCTTTATCTAGGCTGCCACCACTAATGACACTGTCTGTGGCACTCCTGGGSCCAACC 240
 Db 181 ATCTTTATCTAGGCTGCCACCACTAATGACACTGTCTGTGGCACTCCTGGGSCCAACC 240
 Qy 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGAACTGATGTGGCCGACCTGCTTTAT 300
 Db 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGAACTGATGTGGCCGACCTGCTTTAT 300
 Qy 301 GTGCTATGCTCCCTCTCATCATCCTACTACTAGTACAGAGTGGCCCTTCGGGGAG 360
 Db 301 GTGCTATGCTCCCTCTCATCATCCTACTACTAGTACAGAGTGGCCCTTCGGGGAG 360
 Qy 361 CTGCTCTGCAAGCTGTGGCACTTCTTATATCAACCTTTACGGCAGCATCTCTGCTG 420
 Db 361 CTGCTCTGCAAGCTGTGGCACTTCTTATATCAACCTTTACGGCAGCATCTCTGCTG 420
 Qy 421 CTGACCTGCACTCTGTGCAAGTCTCTAGTGTGTGCCACCACTGTTCCTGCTGCC 480
 Db 421 CTGACCTGCACTCTGTGCAAGTCTCTAGTGTGTGCCACCACTGTTCCTGCTGCC 480
 Qy 481 TACCGACCCGAGGATGCTGTGGGACACAGACACCTGGGCTGTGGTCTCTC 540
 Db 481 TACCGACCCGAGGATGCTGTGGGACACAGACACCTGGGCTGTGGTCTCTC 540
 Qy 541 CAGCTGTGCCCACTGTGGGCTTCTCCACACGGACTATCAATGCGCAGATGATCTGG 600
 Db 541 CAGCTGTGCCCACTGTGGGCTTCTCCACACGGACTATCAATGCGCAGATGATCTGG 600
 Qy 601 TATGACATGACAGCAGCAAGATTTTGTGCGCTTTTGTCCATACGCGATGTTCTGACA 660
 Db 601 TATGACATGACAGCAGCAAGATTTTGTGCGCTTTTGTCCATACGCGATGTTCTGACA 660
 Qy 661 TTGCTGCTGCTTTCTTCCCTCTGCTGCTATTTTGGTGTGCTATTCATGATGCTCAGGAG 720
 Db 661 TTGCTGCTGCTTTCTTCCCTCTGCTGCTATTTTGGTGTGCTATTCATGATGCTCAGGAG 720
 Qy 721 CCTGATACGCCAGAGAGACCTCATGAGGACAGGACACAGCCGCGAGCCAGTCCAT 780
 Db 721 CCTGATACGCCAGAGAGACCTCATGAGGACAGGACACAGCCGCGAGCCAGTCCAT 780
 Qy 781 CCGGACCATCTACTGTGTGTGCTGCTCTTCAACCCTCTGTTTTGTGCTCCCTTCATATCAC 840

Db 781 CCGGACCATCTACTGTGTGTGCTCTTCAACCCTCTGTTTTGTGCTTCATATCAC 840
 Qy 841 TCGCTCTCTTACCTTCACCATCTGCTTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 900
 Db 841 TCGCTCTCTTACCTTCACCATCTGCTTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 900
 Qy 901 AGCAGTGTGCTCTACAGATATGAGGCTCTGCTGAGTGTGAGAGCTGCTCAACCC 960
 Db 901 AGCAGTGTGCTCTACAGATATGAGGCTCTGCTGAGTGTGAGAGCTGCTCAACCC 960
 Qy 961 AGTCTGCTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTCAGG 1020
 Db 961 AGTCTGCTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTCAGG 1020
 Qy 1021 CAGAACAAAGTTGGTGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAAACAGATCT 1080
 Db 1021 CAGAACAAAGTTGGTGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAAACAGATCT 1080
 Qy 1081 GG 1082
 Db 1081 GG 1082

RESULT 2

ADB99267

ID ADB99267 standard; cDNA; 1156 BP.

XX ADB99267;

XX 04-DEC-2003 (first entry)

XX Human p2Y2li cDNA.

XX ss; gene; human; P2Y2li; chromosome 3; G protein-coupled receptor; GPCR;
 Class A rhodopsin-like sub-family; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 21..1106

XX /*tag= a

XX /product= "P2Y2li"

XX DE10144044-A1.

XX 27-MAR-2003.

XX 07-SEP-2001; 2001DE-01044044.

XX 07-SEP-2001; 2001DE-01044044.

XX (BRUE/) BRUESS M.

XX (BOEN/) BOENISCH H.

XX (VKUE/) VON KUEGELGEN I.

XX Bruess M, Boenisch H, Von Kuegelgen I;

XX WPI; 2003-364675/35.

XX P-PSDB; ADB99264.

XX New human gene P2Y2li and encoded G protein-coupled receptor, useful for
 treatment and diagnosis of receptor-associated diseases and for drug
 screening.

XX Disclosure; Page 4; 6pp; German.

XX This invention describes the human p2Y2li gene and its 5' - and 3' -
 CC untranslated regions, located on chromosome 3 which is a novel G protein-
 CC coupled receptor (GPCR). The protein encoded by P2Y2li is expressed in
 CC blood cells, testis and embryonal kidney cells and contains potential
 CC sites for phosphorylation by protein kinase C and casein kinase II. It is
 CC a member of the Class A rhodopsin-like sub-family of G protein-coupled

CC receptors and it probably a nucleoside/nucleotide receptor that mediates
CC action of nucleosides/nucleotides or their sugar derivatives. P2Y2L1 and
CC antibodies directed against the encoded protein are useful in diagnosis
CC and treatment (including gene therapy) of diseases associated with
CC abnormal levels of P2Y2L1 expression, in screening assays for modulators,
CC potential therapeutic agents, and to produce transgenic animals, e.g. for
CC identifying diseases associated with abnormal expression of P2Y2L1. This
CC sequence represents the P2Y2L1 gene described in the disclosure of the
CC invention.

XX
SQ Sequence 1156 BP; 243 A; 335 C; 295 G; 283 T; 0 U; 0 Other;
Query Match 100.0%; Score 1082; DB 10; Length 1156;
Best Local Similarity 100.0%; Pred. No. 4e-285;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTGTCCATTTTGTCTTCTCCAGGGGAAGCAGAGGGGAGCGTCTGTGAGCTCTG 60
Db 21 ATGCTGTCCATTTTGTCTTCTCCAGGGGAAGCAGAGGGGAGCGTCTGTGAGCTCTG 80
Qy 61 CTCTGTGAGGAGCGCTCCCGGACATGAGAGGTGAGATGAATACATCACAGGAACAA 120
Db 81 CTCTGTGAGGAGCGCTCCCGGACATGAGAGGTGAGATGAATACATCACAGGAACAA 140
Qy 121 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGCGCTACAGTATC 180
Db 141 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGCGCTACAGTATC 200
Qy 181 ATCTTTATCTAGGCTGCCATTAATGSCACTGTCTTGTGSCACTCTCTGGGGCCAAAC 240
Db 201 ATCTTTATCTAGGCTGCCATTAATGSCACTGTCTTGTGSCACTCTCTGGGGCCAAAC 260
Qy 241 AAGCGTGTGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 300
Db 261 AAGCGTGTGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 320
Qy 301 GTGCTATGTGCTCTCTCATCATCACTACTAGTACAGAGTGAGCGCTTCTGGGGAG 360
Db 321 GTGCTATGTGCTCTCTCATCATCACTACTAGTACAGAGTGAGCGCTTCTGGGGAG 380
Qy 361 CTGCTCTGCAAGTGTGTGCACTTCTTCTTATATCAACCTTTACGCGAGCATCTGCTG 420
Db 381 CTGCTCTGCAAGTGTGTGCACTTCTTCTTATATCAACCTTTACGCGAGCATCTGCTG 440
Qy 421 CTGACCTGCATCTCTGTGCAACCACTCTAGTGTGTGCAACCACTGTGTGCTGCC 480
Db 441 CTGACCTGCATCTCTGTGCAACCACTCTAGTGTGTGCAACCACTGTGTGCTGCC 500
Qy 481 TACGGGACCGAGGATGCTGGTGGCACGACGACCACTGGGCCCTGTGTGCTCTC 540
Db 501 TACGGGACCGAGGATGCTGGTGGCACGACGACCACTGGGCCCTGTGTGCTCTC 560
Qy 541 CAGCTGTGCCCCACACTGGGCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG 600
Db 561 CAGCTGTGCCCCACACTGGGCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG 620
Qy 601 TATGACATGACCGAGCAAGATTTTGATCGGCTTTTTCCTACGGCATAGTCTGACA 660
Db 621 TATGACATGACCGAGCAAGATTTTGTATCGGCTTTTTCCTACGGCATAGTCTGACA 680
Qy 661 TTGTCTGCGCTTCTTCTCCCTCTGTGTCATTTGTGTGCTATTAATCACTGATGTCAAGGAG 720
Db 681 TTGTCTGCGCTTCTTCTCCCTCTGTGTCATTTGTGTGCTATTAATCACTGATGTCAAGGAG 740
Qy 721 CCTGATCAAGCCAGAGGAGAACCTCATAGGACAGGCAACACAGCCGAGCCAGGTCAT 780
Db 741 CCTGATCAAGCCAGAGGAGAACCTCATAGGACAGGCAACACAGCCGAGCCAGGTCAT 800
Qy 781 CCGGACCATCTACTGTGTGTGCGCTCTTACCTCTGTTTGTGCTTCCATATCAC 840
Db 801 CCGGACCATCTACTGTGTGTGCGCTCTTACCTCTGTTTGTGCTTCCATATCAC 860
Qy 841 TCGTCTCTTCTACCTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTGTGCGC 900

Db 861 TCGTCTCTTCTACTCTACCATCTGCTTCTGCTTCTCAGAGCTGCCAGCTCTTGATGC 920
Qy 901 AGCCAGTGTGCGCTTACAGATATGAGAGCCCTCTGTGAGTGTGAGCTGCTCAACCC 960
Db 921 AGCCAGTGTGCGCTTACAGATATGAGAGCCCTCTGTGAGTGTGAGCTGCTCAACCC 980
Qy 961 AGTCTGTACTTCTTCTTCAAGGGGGCAAAATAGATCAGGCTCTCTCCAGAACTGAGG 1020
Db 981 AGTCTGTACTTCTTCTTCAAGGGGGCAAAATAGATCAGGCTCTCTCCAGAACTGAGG 1040
Qy 1021 CAGAACAGTTGGGTGAGCATCCAGCTGGGAGGAGATGCCCGAGGTTGAACAGATCT 1080
Db 1041 CAGAACAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCCGAGGTTGAACAGATCT 1100
Qy 1081 GG 1082
Db 1101 GG 1102

RESULT 3
ADC86166
ID ADC86166 standard; DNA; 1334 BP.
XX
AC ADC86166;
XX AC
XX 01-JAN-2004 (first entry)
XX
XX Human GPCR gene SEQ ID NO:619.
XX
XX ds; gene; human; GPCR;
XX guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX Homo sapiens.
XX
XX EP1270724-A2.
XX
XX 02-JAN-2003.
XX
XX 18-JUN-2002; 2002EP-00013517.
XX
XX 18-JUN-2001; 2001JP-00246789.
XX
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
XX Suwa M, Asai K, Akiyama Y, Aburatani H;
XX
XX WPI: 2003-315783/31.
XX P-PSDB; ADC86167.
XX
XX New polynucleotide, useful for preparing a composition for treating a
XX patient in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor.
XX
XX Claim 1; SEQ ID NO 619; 28pp; English.
XX
XX The invention relates to a novel polynucleotide encoding a guanosine
XX triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX the invention may have a use in gene therapy. The polynucleotide and
XX polypeptide are useful for preparing a composition for treating a patient
XX in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor. The
XX polynucleotide sequences shown in ADC8548-ADC87616 encode GPCR's of the
XX invention.
XX
XX Sequence 1334 BP; 283 A; 368 C; 348 G; 335 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 1082; DB 10; Length 1334;
Best Local Similarity 100.0%; Pred. No. 4.3e-285;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTGTCCATTTTGTCTTCTTCCAGGGGAAGCAGAGGGGAGCGTCTGTGAGCTCTG 60

Db 117 ATGCTGTCATTTTTCCTTTCAGGGGAGCAGAGCGGGAGCCGTGAGGCTCTG 176
 Qy 61 CTCTGAGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGATACATACAGGAGCAA 120
 Db 177 CTCTGAGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGATACATACAGGAGCAA 236
 Qy 121 GGTCTCTGCCAGTTCTCAGAGAGTACAGAGAGTCTACCTCTCCCTGGGCTACATAC 180
 Db 237 GGTCTCTGCCAGTTCTCAGAGAGTACAGAGAGTCTACCTCTCCCTGGGCTACATAC 296
 Qy 181 ATCTTTATCTTAGGGCTGCCAATAATGGCACTGTCTTGTGGCACTCTCGGGGCCAAACC 240
 Db 297 ATCTTTATCTTAGGGCTGCCAATAATGGCACTGTCTTGTGGCACTCTCGGGGCCAAACC 356
 Qy 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGTATGGTGGCGGACCTGCTTTAT 300
 Db 357 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGTATGGTGGCGGACCTGCTTTAT 416
 Qy 301 GTGCTATTGCCCTTCTCATCATCACTACTCTACTAGATCAGAGTGGCCCTTGGGGAG 360
 Db 417 GTGCTATTGCCCTTCTCATCATCACTACTCTACTAGATCAGAGTGGCCCTTGGGGAG 476
 Qy 361 CTGCTCTCAAGCTGTGGCACTTCTCTTATATCAACCTTTACGGCAGCATCTCTGCTG 420
 Db 477 CTGCTCTCAAGCTGTGGCACTTCTCTTATATCAACCTTTACGGCAGCATCTCTGCTG 536
 Qy 421 CTGACCTGCATCTCTGTGCACTTCTCTAGTGTGTGCCACCACTGTGTGCTGCTGCC 480
 Db 537 CTGACCTGCATCTCTGTGCACTTCTCTAGTGTGTGCCACCACTGTGTGCTGCTGCC 596
 Qy 481 TACCGGACCCGAGGAGTCTGTGCTGGCCAGCAGCACCCTGGGCCCTTGGTGGTCTC 540
 Db 597 TACCGGACCCGAGGAGTCTGTGCTGGCCAGCAGCACCCTGGGCCCTTGGTGGTCTC 656
 Qy 541 CAGCTGCTGCCACACTGTGGCTTCTCCACAGGACATGATGATGATGATGATGATG 600
 Db 657 CAGCTGCTGCCACACTGTGGCTTCTCCACAGGACATGATGATGATGATGATGATG 716
 Qy 601 TATGACATGACAGCAGGAGATTTTGTATGCTGCTTTTGTGCTTACGCTAGTCTGACA 660
 Db 717 TATGACATGACAGCAGGAGATTTTGTATGCTGCTTTTGTGCTTACGCTAGTCTGACA 776
 Qy 661 TTGCTGCTGCTTCTTCT 720
 Db 777 TTGCTGCTGCTTCTTCT 836
 Qy 721 CCTGATCAAGCAGGAGAGACCTCATGAGCAGGAGCAACAGCAGCAGCAGGTCCTAT 780
 Db 837 CCTGATCAAGCAGGAGAGACCTCATGAGCAGGAGCAACAGCAGCAGCAGGTCCTAT 896
 Qy 781 CCGGACCATCTACTGTGTGTGCTCTTCAACCTCTGTTTGTGCTTCCATATCAC 840
 Db 897 CCGGACCATCTACTGTGTGTGCTCTTCAACCTCTGTTTGTGCTTCCATATCAC 956
 Qy 841 TCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 900
 Db 957 TCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1016
 Qy 901 AGCCAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 960
 Db 1017 AGCCAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1076
 Qy 961 AGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1020
 Db 1077 AGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1136
 Qy 1021 CAGAACAGTTGGTGTGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1080
 Db 1137 CAGAACAGTTGGTGTGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1196
 Qy 1081 GG 1082

Db 1197 GG 1198
 RESULT 4
 AAS17746
 ID AAS17746 standard; DNA; 3143 BP.
 XX
 AC AAS17746;
 XX 26-FEB-2002 (first entry)
 XX Human genomic clone for P2Y-like G protein-coupled receptor.
 DE Human; ds; P2Y-like G protein-coupled receptor; GPCR; COPD;
 KW chronic obstructive pulmonary disease; nervous system disease;
 KW Parkinson's disease; multiple sclerosis; dementia; stroke; incontinence;
 KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;
 KW bacterial infection; fungal infection; protozoan infection;
 KW viral infection; pain; cancer; anorexia; bulimia; asthma;
 KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;
 KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;
 KW psychotic disorder; neurological disorder; anxiety; schizophrenia;
 KW manic depression; delirium; severe mental retardation; dyskinesia.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 CDS 520..2498
 FT /tag= a
 FT /product= "P2Y-like GPCR"
 XX
 PN WO200185764-A2.
 XX
 PD 15-NOV-2001.
 XX
 PP 09-MAY-2001; 2001WO-EP005244.
 XX
 PR 11-MAY-2000; 2000US-0203582P.
 PR 21-FEB-2001; 2001US-0269857P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Ramakrishnan S;
 XX
 DR WPI; 2002-075242/10.
 DR P-PSDB; AAU11251.
 XX
 PT New polynucleotides for producing P2Y-like G protein-coupled receptors (GPCR) that are used for screening inhibitors or regulators of human P2Y-like GPCR, especially useful for treating pain, cancer or neurological disorders.
 XX
 PS Disclosure; Fig 1; 114pp; English.
 XX
 CC The invention relates to an isolated polynucleotide encoding a P2Y-like G protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or allele, a host cell containing an expression vector comprising the polynucleotide and screening for agents that regulate the GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR polypeptide, which may be employed for screening agents that inhibit or regulate human P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR disorders, particularly COPD (chronic obstructive pulmonary disease), peripheral or central nervous system disease (e.g. Parkinson's disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and many other diseases and disorders listed in the specification), benign prostatic hyperplasia or urinary incontinence. A pharmaceutical composition containing the modulators and/or regulators of P2Y-like GPCR is useful for modulating the activity of a P2Y-like GPCR. In particular, these are useful for treating, preventing or ameliorating infections (e.g. bacterial, fungal, protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma, acute heart failure, hypotension, hypertension, inflammation, diabetes, angina pectoris, myocardial infarction, ulcers, inflammation, allergies,

CC modulating the activity of NOVX, treating or preventing a NOVX-associated
 CC disorder, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX and treating a pathological state
 CC in a mammal comprising administering a polypeptide which is at least 95%
 CC identical to NOVX (or fragment). NOVX and NA may be used in the
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,
 CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and
 CC antagonists may also be used to down regulate expression and activity of
 CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for
 CC detecting the presence of NOVX in samples (e.g. by enzyme linked
 CC immunosorbent assay (ELISA)). The agents and methods may be used in this
 CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy
 CC and/or atherosclerosis. The present sequence encodes a NOVX protein.
 XX
 SQ Sequence 1108 BP; 225 A; 321 C; 285 G; 277 T; 0 U; 0 Other;

Query Match 97.9%; Score 1059; DB 12; Length 1108;
 Best Local Similarity 99.8%; Pred. No. 7.7e-279;
 Matches 1081; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 ATGCTGTCATTTTCCTTCCAGGGGAGCAGAGCGGAGCGGCTGCTGAGCTCTG 60
 Db 12 ATGCTGTCATTTTCCTTCCAGGGGAGCAGAGCGGAGCGGCTGCTGAGCTCTG 71
 Qy 61 CTCCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATCATCACAGGACAA 120
 Db 72 CTCCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATCATCACAGGACAA 131
 Qy 121 GGTCTGCGCAATTTCTCAGAGAAGTCAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
 Db 132 GGTCTGCGCAATTTCTCAGAGAAGTCAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 191
 Qy 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTCTCTGTGGCACTCCCTGGGCCCAACC 240
 Db 192 ATCTTTATCTAGGCTGCCACTAAATGGCACTCTCTGTGGCACTCCCTGGGCCCAACC 251
 Qy 241 AAGCGCTGGAGCTGTGCACCACTATCTGTGTGAACCTGTATGTGGCGCACTGCTTTAT 300
 Db 252 AAGCGCTGGAGCTGTGCACCACTATCTGTGTGAACCTGTATGTGGCGCACTGCTTTAT 311
 Qy 301 GTGCTATGCTTCTCTCATCATCATCTACTACTAGATGACAGGTGGCCCTTGGGGAG 360
 Db 312 GTGCTATGCTTCTCTCATCATCATCTACTACTAGATGACAGGTGGCCCTTGGGGAG 371
 Qy 361 CTGCTCTGCAAGCTGTGTCCTCTCTTATATCACTTTACGCGACATCTCTGCTG 420
 Db 372 CTGCTCTGCAAGCTGTGTCCTCTCTTATATCACTTTACGCGACATCTCTGCTG 431
 Qy 421 CTGACCTGCAATCTGTGCAACAGTCTCTAGGTGTGTGCAACCACTGTGCTGCCCC 480
 Db 432 CTGACCTGCAATCTGTGCAACAGTCTCTAGGTGTGTGCAACCACTGTGCTGCCCC 491
 Qy 481 TACCGGACCGGAGCTGTGCTGGGACACAGCACCACTGGCCCTGTGTGCTC 540
 Db 492 TACCGGACCGGAGCTGTGCTGGGACACAGCACCACTGGCCCTGTGTGCTC 551
 Qy 541 CAGCTGCTGCCACACTGGCCTTCTCCACACGAGCTACATCAATGCCAGATGATCTGG 600
 Db 552 CAGCTGCTGCCACACTGGCCTTCTCCACACGAGCTACATCAATGCCAGATGATCTGG 611
 Qy 601 TATGACATGACCAAGAGAAATTTGATCGGCTTTTTCCTACGGCATAGTTCGACA 660
 Db 612 TATGACATGACCAAGAGAAATTTGATCGGCTTTTTCCTACGGCATAGTTCGACA 671
 Qy 661 TTGCTGCTTCTTCTCCCTGCTGCTATTTTGGTGTGCTATCTACTAGTGTGAGGAG 720
 Db 672 TTGCTGCTTCTTCTCCCTGCTGCTATTTTGGTGTG-TATTACTGATGCTGAGGAG 730
 Qy 721 CCTGATCAGCAGAGAGAACTCATGAGGACAGGCAACAGCCCGAGCCAGTCCAT 780
 Db 731 CCTGATCAGCAGAGAGAACTCATGAGGACAGGCAACAGCCCGAGCCAGTCCAT 790

Qy 781 CGGACCATCTACTGCTGCTGCTTTCACCCCTCTGTTTGGCCTTCCATATCAC 840
 Db 791 CGGACCATCTACTGCTGCTGCTTTCACCCCTCTGTTTGGCCTTCCATATCAC 850
 Qy 841 TCGCTCTCTTACCTCACCATCTGCTTCTCTCTTCTCAGGACTGCCAGCTCTTGTATGGC 900
 Db 851 TCGCTCTCTTACCTCACCATCTGCTTCTCTCTTCTCAGGACTGCCAGCTCTTGTATGGC 910
 Qy 901 AG-CCAGTGTGGCCTTACAGATATGAGAGCTCTGCTGAGTGTGAGAGCTGCTCAACC 959
 Db 911 AGCCAGTGTGGCCTTACAGATATGAGAGCTCTGCTGAGTGTGAGAGCTGCTCAACC 970
 Qy 960 CAGTCTCTTCTTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCTCCAGAACTGAG 1019
 Db 971 CAGTCTCTTCTTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCTCCAGAACTGAG 1030
 Qy 1020 GCAGAACAAAGTTGGTGGAGCATCCAGCTGGGAGGAGAGATGCCAGGTTGAAACAGATC 1079
 Db 1031 GCAGAACAAAGTTGGTGGAGCATCCAGCTGGGAGGAGAGATGCCAGGTTGAAACAGATC 1090
 Qy 1080 TGG 1082
 Db 1091 TGG 1093

RESULT 7

AAS98135
 ID AAS98135 standard; DNA; 1063 BP.
 XX AAS98135;
 XX 12-MAR-2002 (first entry)
 XX Human DNA for potential G protein-coupled receptor #92.
 DE Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;
 KW Alzheimer's disease; amyotrophic lateral sclerosis; asthma;
 KW atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;
 KW chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;
 KW depression; epilepsy; macular degeneration; lymphoma; melanoma;
 KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;
 KW psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;
 KW tuberculosis; cognition disorder; memory disorder; anorexia;
 KW hormonal release disorder; cardiovascular activity disorder;
 KW pain perception disorder; obesity; diabetes; diabetes; obesity; diabetes;
 KW hyperlipidaemia; stroke; gene therapy.
 XX Homo sapiens.
 OS WO200185791-A1.
 FN 15-NOV-2001.
 PD 11-MAY-2001; 2001WO-US015332.
 PF 11-MAY-2000; 2000US-0203217P.
 PR 18-MAY-2000; 2000US-0205945P.
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 XX Brown JP, Miller M, Burmer G, Fabre-Suver C, Pritchard D;
 DR WPI; 2002-065595/09.
 XX

PT Novel G protein-coupled receptor polypeptides including galanin receptor
 PT polypeptides useful for identifying modulators that are useful for
 PT treating Alzheimer's disease, psoriasis, melanoma, multiple sclerosis,
 PT stroke.
 XX

PS Claim 2; Page 98; 144pp; English.

CC The invention relates to an isolated polypeptide encoded by a nucleic
 CC acid molecule that is at least 80% identical to the G protein-coupled

(GPCR) polynucleotides included in the specification. Also included are probes based on the GPCR sequences (including antisense probes), a host cell comprising an expression vector comprising the GPCR sequence, antibodies raised against the polypeptides, and methods of identifying modulators of the polypeptides. The polypeptides are useful for identifying modulator compounds which function as modulators, activators, repressors, agonists or antagonists of the novel GPCR polypeptides including the GAL4 polypeptide. The antibodies and nucleic acid probes as described above can be used to detect the presence of the polypeptides and nucleic acids and are used to diagnose a variety of diseases or disorders in which GPCRs are involved e.g., Alzheimer's disease, anorectic lateral sclerosis, asthma, atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy, chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease, depression, epilepsy, macular degeneration, lymphoma, melanoma, multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease, psoriasis, and many other diseases listed in the specification. The probes and antibodies are also useful for diagnosing cognition and memory disorders, anorexia, hormonal release disorders, cardiovascular activity disorders, pain perception disorders, obesity, diabetes, Alzheimer's disease. Preferably, compounds that decrease or increase the expression of galanin receptor (GAL4) can be used to treat obesity, diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is useful for treating the above mentioned disorders by gene therapy techniques. The present sequence is a novel GPCR polynucleotide of the invention

XX SQ Sequence 1063 BP; 216 A; 315 C; 272 G; 260 T; 0 U; 0 Other;

Query Match 97.1%; Score 1051; DB 6; Length 1063;
 Best Local Similarity 99.9%; Pred. No. 1.2e-276;
 Matches 1062; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 19 CCTTCAGGGGAGCAGAGCGGGAGCGCTGCTGGAGCTCTGCTCTGGAGGAGCGCTCC 78
 DB 1 CCTTCAGGGGAGCAGAGCGGGAGCGCTGCTGGAGCTCTGCTCTGGAGGAGCGCTCC 60
 QY 79 CGGAGATGAGAGAGGTGGATGATATACATACAGGACAAAGTCTCTGCCAGTTCTCA 138
 DB 61 CGGAGATGAGAGAGGTGGATGATATACATACAGGACAAAGTCTCTGCCAGTTCTCA 120
 QY 139 GAGAAGTACAGCAAGTCTACCTCTCCCTGGCTACAGTATATCTTTATCTTAGGGCTG 198
 DB 121 GAGAAGTACAGCAAGTCTACCTCTCCCTGGCTACAGTATATCTTTATCTTAGGGCTG 180
 QY 199 CCACTAATGGCACTCTCTGTCGCACTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCC 258
 DB 181 CCACTAATGGCACTCTCTGTCGCACTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCC 240
 QY 259 ACCACCTATCTGCTGAACCTGTGTCGCACTCTGTTTATCTGCTATTGCCCTTCTCTC 318
 DB 241 ACCACCTATCTGCTGAACCTGTGTCGCACTCTGTTTATCTGCTATTGCCCTTCTCTC 300
 QY 319 ATCATCACTTACTACTAGATGACAGTGCGCCCTTCTGGGAGCTGCTTGCAGCTGGTG 378
 DB 301 ATCATCACTTACTACTAGATGACAGTGCGCCCTTCTGGGAGCTGCTTGCAGCTGGTG 360
 QY 379 CACTTCTGTTCTATATCAACCTTTACGGCAGATCTGCTGCTGACCTGCATCTCTGTG 438
 DB 361 CACTTCTGTTCTATATCAACCTTTACGGCAGATCTGCTGCTGACCTGCATCTCTGTG 420
 QY 439 CACAGTCTTCTAGTGTGTCACCCACTGTTGTCGCTGCCCTTACCGGACCCGAGCAT 498
 DB 421 CACAGTCTTCTAGTGTGTCACCCACTGTTGTCGCTGCCCTTACCGGACCCGAGCAT 480
 QY 499 GCGTGGTGGGACCAACCACTGGGCGCTGGTGTCTCCAGTGTGCTGCCACACTG 558
 DB 481 GCGTGGTGGGACCAACCACTGGGCGCTGGTGTCTCCAGTGTGCTGCCACACTG 540
 QY 559 GCCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGTGATGACATGACCGCAA 618
 DB 541 GCCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGTGATGACATGACCGCAA 600

QY 619 GAGAATTTGATCGGCTTTTGGCTTACGGCATAGTTCTGACATGTCTGGCTTCTTTTCC 678
 DB 601 GAGAATTTGATCGGCTTTTGGCTTACGGCATAGTTCTGACATGTCTGGCTTCTTTTCC 660
 QY 679 CTCCTTGGTCAATTTTGGTGTCTATTCTGATGCTCAGGAGCTGATCAAGCCAGAGGA 738
 DB 661 CTCCTTGGTCAATTTTGGTGTCTATTCTGATGCTCAGGAGCTGATCAAGCCAGAGGA 720
 QY 739 GAACTCTATGAGGACAGGCAACAGCCCGGAGCCAGGTCCATCCGACCATCTACTGGT 798
 DB 721 GAACTCTATGAGGACAGGCAACAGCCCGGAGCCAGGTCCATCCGACCATCTACTGGT 780
 QY 799 GTGTGGCTCTTACCCCTCTGTTTGTGGCTTCCATATCATCTGCTCTTCTACCTCAC 858
 DB 781 GTGTGGCTCTTACCCCTCTGTTTGTGGCTTCCATATCATCTGCTCTTCTACCTCAC 840
 QY 859 CATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGSCAG-CCAGTGTGGCCTACA 917
 DB 841 CATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGSCAGCCAGTGTGGCCTACA 900
 QY 918 AGATATGAGGCGCTCTGCTGATGAGTGTGAGCTGCTCCCAACCCAGTCTCTACTTCTT 977
 DB 901 AGATATGAGGCGCTCTGCTGATGAGTGTGAGCTGCTCCCAACCCAGTCTCTACTTCTT 960
 QY 978 CAAGGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTGAGGAGCAACAAAGTTGGGTGA 1037
 DB 961 CAAGGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTGAGGAGCAACAAAGTTGGGTGA 1020
 QY 1038 GCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT 1080
 DB 1021 GCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT 1063

RESULT 8
 ADC25999 standard; DNA; 1017 BP.
 XX ID ADC25999 standard; DNA; 1017 BP.
 AC ADC25999;
 DX 18-DEC-2003 (first entry)
 XX Human purinergic receptor P2Y-related GPCRx6 DNA.
 DE virucide; fungicide; antibacterial; cytostatic; analgesic; antidiabetic;
 KW anorectic; cardiant; hypotensive; osteopathic; anti-anginal;
 KW antiarteriosclerotic; cerebroprotective; anti-ulcer; antiallergic;
 KW neurotropic; neuroprotective; antiparkinsonian; G-protein coupled receptor;
 KW GPCR; viral; fungal; bacterial infection; immune-related disorder;
 KW cancer; pain; diabetes; obesity; anorexia; acute heart failure;
 KW hypertension; osteoporosis; angina pectoris; atherosclerosis; stroke;
 KW ulcer; allergy; psychotic neurological disorder; schizophrenia; dementia;
 KW degenerative disease; Parkinson's; Alzheimer's; dyskinesia; Huntington's;
 KW human; GPCRx6; purinergic receptor P2Y; ds; gene.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 CDS 1..1017
 /tag= a
 /product= "Human purinergic receptor P2Y-related GPCRx6
 protein - both "original" and "alternative" versions"
 /transl_except= (pos:184..186, aa:Pro)
 /transl_except= (pos:727..729, aa:Xaa)
 /transl_except= (pos:808..810, aa:Pro)
 /note= "Xaa = Unknown; translation exceptions at 184 and
 727 are present within "original" protein CDS whilst that
 at 808 is present within the "alternative" protein CDS"
 XX US2003088080-A1.
 XX 08-MAY-2003.
 XX 21-JUN-2001; 2001US-00885453.
 PF

Db	481	TACCGGACCGGAGCATGCTGGCTGGGACACGACCACTGGGCGCTGGTGGCTTC	540
Qy	541	CAGCTGCTGCCACACTGGCCTTCTCCACACGAGACTACATCAATGGCCAGATGATCTGG	600
Db	541	CAGCTGCTGCCACACTGGCCTTCTCCACACGAGACTACATCAATGGCCAGATGATCTGG	600
Qy	601	TATGACATGACAGCCAGAGAAATTTTGATCGGCTTTTTCCTACGGCATAGTTCTTGACA	660
Db	601	TATGACATGACAGCCAGAGAAATTTTGATCGGCTTTTTCCTACGGCATAGTTCTTGACA	660
Qy	661	TTGTCTGGCTTTCTTTCCCTCCTTGGTCAATTTTGGTGTGCTATTCACTGATGCTCAGGAG	720
Db	661	TTGTCTGGCTTT- TTTCCCTCCTTGGTCAATTTTGGTGTGCTATTCACTGATGCTCAGGAG	719
Qy	721	CTGTGATCAGCCAGAGGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGTCCTAT	780
Db	720	CTGTGATCAGCCAGAGGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGTCCTAT	779
Qy	781	CCGGACCACTCCTACTGGTGTGTGGCCTCTTACACCTCTGTTTTGTGCGCTTCCATATCAC	840
Db	780	CCGGACCACTCCTACTGGTGTGTGGCCTCTTACACCTCTGTTTTGTGCGCTTCCATATCAC	839
Qy	841	TGCTTCCTTACTACCACTCGCTTCTGCTTTCTCAGGACTGCGAGCTCTTGATGGC	900
Db	840	TGCTTCCTTACTACCACTCGCTTCTGCTTTCTCAGGACTGCGAGCTCTTGATGGC	899
Qy	901	ASCCAGTGTGGCTACAGATATGAGGCGCTCTGGTGAGTGTGAGCAGCTGCTCAACCC	960
Db	900	ASCCAGTGTGGCTACAGATATGAGGCGCTCTGGTGAGTGTGAGCAGCTGCTCAACCC	959
Qy	961	AGTCCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTTCAGAAACTGA	1018
Db	960	AGTCCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTTCAGAAACTGA	1017
RESULT 9			
AAD61648			
ID AAD61648 standard; DNA; 1017 BP.			
XX AAD61648;			
XX AC AAD61648;			
DT 15-JAN-2004 (first entry)			
XX Human GPCR6 DNA.			
XX Human; G-protein coupled receptor; GPCR; infection; neoplastic process;			
XX inflammation; myocardial infarction; atherosclerosis; angina pectoris;			
XX hypertension; osteoporosis; antibacterial; cytostatic; fungicide; pain;			
XX diabetes; cancer; virucide; analgesic; cardiant; gene; ds.			
XX Homo sapiens.			
XX OS			
XX Key Location/Qualifiers			
XX CDS 1..1017			
XX FT /tag= a			
XX FT /product= "Human GPCR protein"			
XX US2003108986-A1.			
XX PN			
XX PD 12-JUN-2003.			
XX PP 20-FEB-2002; 2002US-00079384.			
XX PR 21-JUN-2001; 2001US-00885453.			
XX (EURO-) EUROSREEN SA.			
XX Communi D, Lannoy V, Brezillon S, Detheux M, Parmentier M;			
XX Govaerts C;			
XX WPI; 2003-810852/76.			
XX DR P-PSDB; ABW00804.			

Query Match		92.7%; Score 1002.8; DB 6; Length 1020;
Best Local Similarity		99.6%; Pred. No. 1.7e-263;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;		
Qy	1	ATGCTGTCCATTTTCTCTCCCTCCAGGGAGACAGAGCGGGAGCCCTCGTGGAGCTCTG 60
Db	1	ATGCTGTCCATTTTCTCTCCCTCCAGGGAGACAGAGCGGGAGCCCTCGTGGAGCTCTG 60
Qy	61	CTCCTGGAGGAGCTCCCGGACATGGAGAGGTGGACATGATATACACAGGAACAA 120
Db	61	CTCCTGGAGGAGCTCCCGGACATGGAGAGGTGGACATGATATACACAGGAACAA 120
Qy	121	GGTCTCTGCCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCTCCCTGGCTACAGTATC 180
Db	121	GGTCTCTGCCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCTCCCTGGCTACAGTATC 180
Qy	181	ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTGTGGCACTCTCTGGGGCCAAAC 240
Db	181	ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTGTGGCACTCTCTGGGGCCAAAC 240
Qy	241	AAGCGCTGGAGCTGTGCCACCACTATCTGTGAACCTGATGTGGCGGACCTGCTTTAT 300
Db	241	AAGCGCTGGAGCTGTGCCACCACTATCTGTGAACCTGATGTGGCGGACCTGCTTTAT 300
Qy	301	GTGCTATTGCCCTTCTCATCATCCTACTCTACTAGATGACAGGTGGCCCTTCGGGGAG 360
Db	301	GTGCTATTGCCCTTCTCATCATCCTACTCTACTAGATGACAGGTGGCCCTTCGGGGAG 360
Qy	361	CTGCTCTGACAGCTGTGCACTCTCTCTATATCAACCTTTACGCGACATCTCTGCTG 420
Db	361	CTGCTCTGACAGCTGTGCACTCTCTCTATATCAACCTTTACGCGACATCTCTGCTG 420
Qy	421	CTGACCTGCACTCTCTGTCACAGTCTCTAGTGTGTGCCACCACTGTCTTCGCTGCC 480
Db	421	CTGACCTGCACTCTCTGTCACAGTCTCTAGTGTGTGCCACCACTGTCTTCGCTGCC 480
Qy	481	TACGGACCCGAGGATGCTGCTGGGACACAGCACCACTGGGCGCTTGTGGTCTCTC 540
Db	481	TACGGACCCGAGGATGCTGCTGGGACACAGCACCACTGGGCGCTTGTGGTCTCTC 540
Qy	541	CAGCTGTGCTCCACACTGGGCTTCTCCACACAGCACTACATCAATGCCAGATGATCTGG 600
Db	541	CAGCTGTGCTCCACACTGGGCTTCTCCACACAGCACTACATCAATGCCAGATGATCTGG 600
Qy	601	TATGACATGACAGCCAGAGAAATTTGATCGGCTTTTTCCTACGCACTAGTCTTGACA 660
Db	601	TATGACATGACAGCCAGAGAAATTTGATCGGCTTTTTCCTACGCACTAGTCTTGACA 660
Qy	661	TTGCTGTGCTTTCTTTT--CCCTCTGCTGCTATTTTGGTGTGCTATTCACATGATGTCAGG 718
Db	661	TTGCTGTGCTTTCTTTT--CCCTCTGCTGCTATTTTGGTGTGCTATTCACATGATGTCAGG 720
Qy	719	AGCCTGATCAAGCCAGAGGAACCTCATGAGGACAGCAACAGCCCGAGCCAGCTCC 778
Db	721	AGCCTGATCAAGCCAGAGGAACCTCATGAGGACAGCAACAGCCCGAGCCAGCTCC 780
Qy	779	ATCCGGACCATCTACTGTTGTGGCTCTTCAACCTCTGTTTGTGCGCTCCATATC 838
Db	781	ATCCGGACCATCTACTGTTGTGGCTCTTCAACCTCTGTTTGTGCGCTCCATATC 840
Qy	839	ACTGCTCTCTTACTACCTACCATCTGCTCTCTGCTTTCTCAGGACTGCCAGCTCTTGATG 898
Db	841	ACTGCTCTCTTACTACCTACCATCTGCTCTCTGCTTTCTCAGGACTGCCAGCTCTTGATG 900
Qy	899	GCACCACTGTGGCTCAGAGATGAGGCGCTCTGTGTGATGTGACAGCTGCTCTCAAC 958
Db	901	GCACCACTGTGGCTCAGAGATGAGGCGCTCTGTGTGATGTGACAGCTGCTCTCAAC 960
Qy	959	CCAGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCTCCAGAACTGA 1018
Db	961	CCAGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCTCCAGAACTGA 1020

RESULT 11	
AAD29667	
ID	AAD29667 standard; cDNA; 1076 BP.
XX	
AC	AAD29667;
XX	
DT	17-MAY-2002 (first entry)
XX	
DE	Human G-protein coupled receptor (GCRC-1) cDNA.
XX	
KW	Human; G-protein coupled receptor; GCRC-1; cell proliferative disorder;
KW	neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;
KW	metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;
KW	Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
KW	atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;
KW	osteoporosis; Crhn's disease; acquired immunodeficiency syndrome; AIDS;
KW	anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
KW	transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;
KW	neurotic; neuroprotective; cardiant; immunosuppressive; anorectic;
KW	viricide; gene; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	57..1076
FT	/*tag= a
XX	/product= "Human GCRC-1"
XX	
PN	WO200210387-A2.
XX	
PD	07-FEB-2002.
XX	
PF	25-JUL-2001; 2001WO-US023433.
XX	
PR	27-JUL-2000; 2000US-0221478P.
PR	03-AUG-2000; 2000US-023268P.
PR	21-AUG-2000; 2000US-027054P.
PR	08-SEP-2000; 2000US-023112P.
PR	13-SEP-2000; 2000US-0232243P.
PR	15-SEP-2000; 2000US-0232691P.
PR	22-SEP-2000; 2000US-0235145P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;
PI	Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walia NK, Hafalia AJA;
PI	Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney JL, Graul RC;
PI	Warren BA, Lee EA, Ding L;
XX	
DR	WPI: 2002-188744/24.
DR	P-PSDB; AAE18540.
XX	
PT	New human G-protein coupled receptor polypeptide for diagnosis,
PT	prevention and treatment of cell proliferative, neurological,
PT	cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic
PT	disorders.
XX	
PS	Claim 5; Page 137-138; 150pp; English.
XX	
CC	The invention relates to novel human G-protein coupled receptors (GCRC)
CC	and their encoding polynucleotides. GCRC is useful as an immunogen for
CC	preparing monoclonal and polyclonal antibodies. GCRC is useful for
CC	diagnosing, treating and preventing a cell proliferative disorder (e.g.,
CC	hepatitis, psoriasis, cancer); a neurological disorder (e.g., epilepsy,
CC	Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's
CC	disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,
CC	myocardial infarction), gastrointestinal disorder (e.g., gastritis,
CC	cirrhosis, Crhn's disease), an autoimmune/inflammatory disorder (e.g.,
CC	acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,
CC	rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,
CC	osteoporosis), and viral infections. GCRC is useful in a number of drug
CC	screening techniques, and to analyse the proteome of a tissue or cell

CC type. GCREC is useful for creating knockin humanised animals or
 CC transgenic animals to model human diseases, in somatic or germline gene
 CC therapy, to generate a transcript image of a tissue or cell type, for
 CC detecting differences in the chromosomal location due to translocation,
 CC inversion, etc., among normal, carrier or affected individuals, and as
 CC hybridization probes for mapping naturally occurring genomic sequences.
 CC GCREC is useful in Southern or northern analysis, dot blot or other
 CC membrane-based technologies, in PCR technologies, in dipstick, pin,
 CC microformat enzyme linked immunosorbant (ELISA)-like assays, and in
 CC microarrays utilising fluids or tissues from patients to detect altered
 CC GCREC expression. The present sequence is human GCREC-1 cDNA
 XX
 SQ Sequence 1076 BP; 211 A; 318 C; 274 G; 273 T; 0 U; 0 Other;

Query Match 92.7%; Score 1002.8; DB 6; Length 1076;
 Best Local Similarity 99.6%; Pred. No. 1-7e-263;
 Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 1 ATGCTGTCCATTTTGTCTCTTCCAGGGGAAGCAGAGCGGAGCGTCTGTGAGTCTTG 60
 Db 57 ATGCTGTCCATTTTGTCTCTTCCAGGGGAAGCAGAGCGGAGCGTCTGTGAGTCTTG 116
 Qy 61 CTCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATGATATACATCAGGAGCAAA 120
 Db 117 CTCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATGATATACATCAGGAGCAAA 176
 Qy 121 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTTACCTCTCCCTGGCTTACAGTATC 180
 Db 177 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTTACCTCTCCCTGGCTTACAGTATC 236
 Qy 181 ATCTTTATCTAGGCTGCGACCAATGAGCACTGTCTTGTGGCACTCTCTGGGCGCAAAAC 240
 Db 237 ATCTTTATCTAGGCTGCGACCAATGAGCACTGTCTTGTGGCACTCTCTGGGCGCAAAAC 296
 Qy 241 AAGCGCTGGAGCTGTGCGACCACTATCTGGTGAACCTGATGTGGCGACCTGCTTTAT 300
 Db 297 AAGCGCTGGAGCTGTGCGACCACTATCTGGTGAACCTGATGTGGCGACCTGCTTTAT 356
 Qy 301 GTGCTATTGGCCCTTCTCATCATCACTATCTAGTAGACAGGTGGCCCTTCTGGGGAG 360
 Db 357 GTGCTATTGGCCCTTCTCATCATCACTATCTAGTAGACAGGTGGCCCTTCTGGGGAG 416
 Qy 361 CTGCTCTGCAAGCTGTGCACTCTCTCTATATCAACCTTTACGGCAGCATCTGCTG 420
 Db 417 CTGCTCTGCAAGCTGTGCACTCTCTCTATATCAACCTTTACGGCAGCATCTGCTG 476
 Qy 421 CTGACCTGCATCTCTGTGCACTCTCTAGGTGTGTGCCACCACTGTGTCTGCTGCC 480
 Db 477 CTGACCTGCATCTCTGTGCACTCTCTAGGTGTGTGCCACCACTGTGTCTGCTGCC 536
 Qy 481 TACGGACCCGACGATGCTGTGCGACCACTGTGCGACCACTGTGCGCTGCTCTC 540
 Db 537 TACGGACCCGACGATGCTGTGCGACCACTGTGCGACCACTGTGCGCTGCTCTC 596
 Qy 541 CAGCTGTGCGACCACTGTGCGACCACTGTGCGACCACTGTGCGCTGCTCTC 600
 Db 597 CAGCTGTGCGACCACTGTGCGACCACTGTGCGACCACTGTGCGCTGCTCTC 656
 Qy 601 TATGACATGACACGCAAGAGAAATTTGATCGGCTTTTGTGCTACGGCATAGTCTGACA 660
 Db 657 TATGACATGACACGCAAGAGAAATTTGATCGGCTTTTGTGCTACGGCATAGTCTGACA 716
 Qy 661 TTGTCTGGCTTCTCTT--CCCTCTCTGCTATTTTGTGTGCTATTCACTGATGCTCAGG 718
 Db 717 TTGTCTGGCTTCTCTTCCCTCTCTTGTGCTATTTGTGCTATTCACTGATGCTCAGG 776
 Qy 719 AGCTGTATCAAGCCAGGAGAACTCATGAGGACAGGCAACACAGCCGAGCGAGTCC 778
 Db 777 AGCTGTATCAAGCCAGGAGAACTCATGAGGACAGGCAACACAGCCGAGCGAGTCC 836
 Qy 779 ATCCGACCATCTTCTGCTGTGCTCTTCTACCTCTGTTTGTGCTTCCATATC 838
 Db 837 ATCCGACCATCTTCTGCTGTGCTCTTCTACCTCTGTTTGTGCTTCCATATC 896

Qy 839 ACTGCTCTTCTTACTCTACCATCTGCTTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATG 898
 Db 897 ACTGCTCTTCTTACTCTACCATCTGCTTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATG 956
 Qy 899 GGAGCAGTGTGGCTTCAAGATATGAGAGGCTCTGCTGAGTGTGAGCAGTCTGCTCAAC 958
 Db 957 GCACCCAGTGTGGCTTCAAGATATGAGAGGCTCTGCTGAGTGTGAGCAGTCTGCTCAAC 1016
 Qy 959 CCAGTCCCTGTACTTCTTCTTCAAGGGGGGCAAAATAGAGTCTGCTTCCAGAACTGA 1018
 Db 1017 CCAGTCCCTGTACTTCTTCTTCAAGGGGGGCAAAATAGAGTCTGCTTCCAGAACTGA 1076

RESULT 12

AD030395 standard; cDNA; 1076 BP.

XX AD030395;

XX 29-JUL-2004 (first entry)

XX Human GPCR P2Y3L polynucleotide, SEQ ID NO:1498.

XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; antineoplastic;
 KW cystostatic; antiinflammatory; vasotropic; antidiarrhoeic; antiarrhythmic;
 KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
 KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
 KW dermatological; antitumor; antithyroid; antiallergic; anorectic;
 KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KW Gene; ss.

XX Homo sapiens.

XX WO2004040000-A2.

XX 13-MAY-2004.

XX 09-SEP-2003; 2003WO-US028226.

XX 09-SEP-2002; 2002US-0409303P.

XX 09-APR-2003; 2003US-0461329P.

XX (PRIM-) PRIMAL INC.

XX Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
 PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;

XX WPI; 2004-390329/36.

XX P-PSDB; AD030394.

XX Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.

XX Claim 151; SEQ ID NO 1498; 542pp; English.

XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a

657	TATGACATGACCGACGACAGAGAAATTTTGATCGCGCTTTTTTGGCTACTACGGCATAGTTCGTACA	716
661	TTGCTGGCTTTCTTT--CCCTCCTTGTCATATTTGGTGTGCTATTCATGTAGTGCAG	718
717	TTGCTGGCTTTCTATTCGCCCTCCTTGTCATATTTGGTGTGCTATTCATGTAGTGCAG	776
719	ASCTGTATCAAGCCAGAGAGAGAACTCATGAGCAGAGGCAACACACAGCCCGAGCCAGCGTCC	778
777	ASCTGTATCAAGCCAGAGAGAGAACTCATGAGCAGAGGCAACACACAGCCCGAGCCAGCGTCC	836
779	ATCCGGAACATCCTACTGTGTGGCGCTCTTCAACCTCTGTTTTGTGCGCTTCCATATC	838
837	ATCCGGAACATCCTACTGTGTGTGGCGCTCTTCAACCTCTGTTTTGTGCGCTTCCATATC	896
839	ACTCGCTCCTTACTCCTCACCATCTGCTTTCTTGCTTTCTCAGCACTGCCAGCTCTTTGATG	898
897	ACTCGCTCCTTACTCCTCACCATCTGCTTTCTTGCTTTCTCAGCACTGCCAGCTCTTTGATG	956
899	GCAGCCAGTGTGGCCTACAGATATGGAGGCGCTCTGGTGAGTGTGAGCAGCTGCCTCAAC	958
957	GCACCCAGTGTGGCCTACAGATATGGAGGCGCTCTGGTGAGTGTGAGCAGCTGCCTCAAC	1016
959	CCAGTCTCTGTACTTTCTTTTCAAGGGGGGCAAAATAGAGTCAAGGTCTCTCCAGAACTGA	1018
1017	CCAGTCTCTGTACTTTCTTTTCAAGGGGGGCAAAATAGAGTCAAGGTCTCTCCAGAACTGA	1076

601 TATGACATGACCAGCCCAAGAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCI

CC	encod
CC	prote

CC present sequence is exon 1 of the P2Y-like GPCR of the invention

XX Sequence 850 BP; 165 A; 262 C; 201 G; 222 T; 0 U; 0 Other;

Query Match 78.6%; Score 850; DB 6; Length 850;
Best Local Similarity 100.0%; Pred. No. 9.2e-222;
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 ATGAGAGAGTGGACATGATACATACATACAGGAAACAGAGTCTGCGCAGTTCTCAGAGAAG 144
Db 1 ATGAGAGAGTGGACATGATACATACATACAGGAAACAGAGTCTGCGCAGTTCTCAGAGAAG 60
QY 145 TACAAGCAAGTCTACCTCTCCCTGCGCTACAGTATCATCTTTATCTTAGGGCTGCCACTA 204
Db 61 TACAAGCAAGTCTACCTCTCCCTGCGCTACAGTATCATCTTTATCTTAGGGCTGCCACTA 120
QY 205 AATGGCACTGTCTGTGGCACTCTCTGGGGCAACCAAGCGCTGGAGCTGTGCCACCAAC 264
Db 121 AATGGCACTGTCTGTGGCACTCTCTGGGGCAACCAAGCGCTGGAGCTGTGCCACCAAC 180
QY 265 TATCTGGTGAACCTGTGGTGGCGACCTCTTTATGTGTATGTGCTTCCCTTCATCATC 324
Db 181 TATCTGGTGAACCTGTGGTGGCGACCTCTTTATGTGTATGTGCTTCCCTTCATCATC 240
QY 325 ACCTACTCAGTATGATGAGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGTGTGCACTTC 384
Db 241 ACCTACTCAGTATGATGAGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGTGTGCACTTC 300
QY 385 CTGTTCTATATCAACCTTTACGGCAGATCTCTGCTGCTGCAAGCTGTGTGCACTTC 444
Db 301 CTGTTCTATATCAACCTTTACGGCAGATCTCTGCTGCTGCAAGCTGTGTGCACTTC 360
QY 445 TTCCTAGTGTGTGCGCACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
Db 361 TTCCTAGTGTGTGCGCACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 505 CTGGGACACAGCACACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564
Db 421 CTGGGACACAGCACACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 565 TCCACACAGGACTACATCAATGGCCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 624
Db 481 TCCACACAGGACTACATCAATGGCCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 625 TTTGATCGGCTTTTGGCTACGGCATAGTTCTGCAATGTCTGGCTTTCTTCCCTCCCT 684
Db 541 TTTGATCGGCTTTTGGCTACGGCATAGTTCTGCAATGTCTGGCTTTCTTCCCTCCCT 600
QY 685 GGTCAATTTGGTGTCTTACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
Db 601 GGTCAATTTGGTGTCTTACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 745 CATGAGACAGGCAACACAGCCCGAGCAGTCCATCCGACCATCTCTACTGTGTGTGG 804
Db 661 CATGAGACAGGCAACACAGCCCGAGCAGTCCATCCGACCATCTCTACTGTGTGTGG 720
QY 805 CCTCTTCACTCTGTTTGGCTTCCATATCATCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 864
Db 721 CCTCTTCACTCTGTTTGGCTTCCATATCATCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 780
QY 865 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGTGAGGACGAGTGTGGCTTACAGATATG 924
Db 781 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGTGAGGACGAGTGTGGCTTACAGATATG 840
QY 925 GAGGCTCTG 934
Db 841 GAGGCTCTG 850

RESULT 15

ADO30397 standard; cDNA; 906 BP.

XX

AC ADO30397;
XX 29-JUL-2004 (first entry)
DT
XX Mouse GPCR P2Y3L polynucleotide, SEQ ID NO:1500.
DE
XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antitumor;
KW cytostatic; antiinflammatory; vasotropic; antiarrhythmic;
KW CNS; central nervous system; respiratory; antiarrhythmic; antidiabetic;
KW virucide; hepatotropic; antibacterial; antianemic; antiseborrheic;
KW dermatological; antitumor; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; mouse;
KW murine; gene; ss.
XX Mus musculus.
OS
XX WO2004040000-A2.
XX 13-MAY-2004.
XX
XX 09-SEP-2003; 2003WO-US028226.
XX
XX 09-SEP-2002; 2002US-0409303P.
XX 09-APR-2003; 2003US-0461329P.
XX (PRIM-) PRIMAL INC.
XX
XX Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
PI
XX WPI; 2004-390329/36.
XX P-PSDB; ADO30396.
XX
XX Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
XX Claim 151; SEQ ID NO 1500; 542pp; English.
XX

The invention relates to human and mouse G protein-coupled receptors (GPCRs) and nucleic acids encoding them. The invention also relates to sequences at least 90% identical to the GPCR proteins and nucleic acids of the invention; methods of treating, preventing or diagnosing diseases associated with GPCRs of the invention; methods of screening for compounds useful in the treatment of GPCR-related diseases; a transgenic mouse comprising a GPCR gene of the invention; a mouse comprising a mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived from the transgenic mice; kits comprising several mice, each of which has a mutation in a different GPCR gene of the invention; and kits comprising probes which hybridize to GPCR polynucleotides of the invention. The invention further discloses variants of the GPCR polypeptides and vectors comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may be used in the diagnosis, treatment or prevention of a wide variety of diseases including neurological disorders (e.g., Alzheimer's disease, depression, diabetic neuropathy, Parkinson's disease or schizophrenia; disorders of the adrenal gland; disorders of the colon or intestine (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or myocardial infarction); muscular disorders; blood disorders (e.g., anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g., obesity, enzyme deficiency-related diseases or vitamin deficiency-related diseases); and disorders of the kidney, liver, lung, breast, ovary, CC

CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences.

XX
SQ Sequence 906 BP; 173 A; 276 C; 216 G; 241 T; 0 U; 0 Other;

Query Match 56.0%; Score 606.2; DB 12; Length 906;
Best Local Similarity 80.9%; Pred. No. 4.1e-155;
Matches 732; Conservative 0; Mismatches 168; Indels 5; Gaps 2;

Qy 85 ATGGAGAAGGTGACATGAATACATACAGGAACAAGGTCTTCCAGTTCCTCAGAGAAG 144
Db |||||
Qy 1 ATGGAGAAGGTAGATATGAATGTCTCAAGGGTCTGGGGCCCTGCCACTTCTCAGAGAGC 60
Db |||||
Qy 145 TACAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTTATCTAGGGCTGCCACTA 204
Db |||||
Qy 61 TACAGCAAGTCTACCTATCATTAACCTACAGTCTCATCTTTATGTAGGATGCCCTC 120
Db |||||
Qy 205 AATGGCACTCTTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACACC 264
Db |||||
Qy 121 AATGGGGCTCTCTATGGCTCTCTGTGGCCAAACCAAGCGCTGGAGCTGTGCCACATC 180
Db |||||
Qy 265 TATCTGGTGAACCTGATGTGGCGGACCTGCTTTATGTGCTA---TTGCCCTTCTCTATC 321
Db |||||
Qy 181 TACCTGATGAACCTGGGGTGGCTGACCTCTTTATGTAGCTTATGTAGCTTCTGCTCAIC 240
Db |||||
Qy 322 ATCACTTACTCTAGTATGACAGTGGCCCTTGGGGAGCTGCTCTGCAAGCTGGTGCAC 381
Db |||||
Qy 241 ATCACTTACGGGTGGGTGACACCTGGCTCTTTGGGGAGATGCTGTGCAGTTGGTGGG 300
Db |||||
Qy 382 TTCTGTTCATATCAACCTTTAGCGCAGATCTCTGTCTGACCTGCATCTCTGTGCAC 441
Db |||||
Qy 301 TTCTGTTCATATCAACCTTTAGCGCAGATCTCTGTCTGACCTGCATCTCTGTGCAC 360
Db |||||
Qy 442 CAGTTCTAGTGTGTCGCCACCACTGTGTGCTGCTGCCCTACCGGACCCCGCAGGCATGCC 501
Db |||||
Qy 361 CGTTCTCTAGAGTGTGCCACCTCTGAGCTCACTGTCTCTACCGACTCACCGGCACGCC 420
Db |||||
Qy 502 TGGCTGGGCAACGACCACTGGGGCCCTGGTGGTCTCTCAGCTGTGGCCCACTGGCC 561
Db |||||
Qy 421 TGGCAGGAGCTGTGTGCCACCTGGACCCCTGGTAGTCTTCCAATTGTGCGGACTCTGGTC 480
Db |||||
Qy 562 TTCTCCACAGGACTATCAATGAGCCAGATGATGATGATGATGATGATGATGATGATGATG 621
Db |||||
Qy 481 TTCTCCACAGGACTATGTAATGAGCCAGTATGATGATGATGATGATGATGATGATGATG 540
Db |||||
Qy 622 AATTTTGTATGGCTTTTGTGCTACGGCATAGTTCTGACATTTGTGCTT - TCTTTCC 679
Db |||||
Qy 541 CAGTTTGTATCAATTCTTCGGCTACAGCATAGTACTGACGTTATCTGGGTTTGTCTCCCC 600
Db |||||
Qy 680 TCCTTGTTCATTTTGGTGTGCTATTCATGATGATGATGATGATGATGATGATGATGATGATG 739
Db |||||
Qy 601 TCCTTGTATCATTTCTGGTGTGCTACTCTCTGATGATGATGATGATGATGATGATGATGATG 660
Db |||||
Qy 740 AACCTCATAGGAGCAGGCAACACAGCCCGAGCCAGGTCCATCCGGACCATCCTACTGGTG 799
Db |||||
Qy 561 GACCTCCCAACGACAGGCCACACAGCCGTCGCAAGTCTATTGGACCATCCTCTCTGGTG 720
Db |||||
Qy 800 TGTGGCTCTTCAACCTCTGTGTTTGTGGCCCTTCCATATCACTGCTCTCTCTCTCTCTCTC 859
Db |||||
Qy 721 TGTGGCTCTTTCACACTCTGTTTGTGGCCCTTCCACATTTGCCCGAATCTTCTACTTGTATC 780
Db |||||
Qy 860 ATCTGCTTTCGCTTCTCAGGACTGCCAGCTCTTGTATGGCAGCAGTGTGGCCTACAAG 919
Db |||||
Qy 781 ATCCGCTTTCCTGAAGTCAAGGACTGCGCAGCTCTTGTGGGAGCCAGCATGGCCTACAAG 840
Db |||||
Qy 920 ATATGGAGCCCTCTGGTGAAGTGTGAGCAGTGCCTCAACCCAGTCTCTGTACTTTCTTCA 979
Db |||||
Qy 841 GTCTGTAGGCTCTGTTAAGCATGAGCAGTGTGCTCAATCCAGTCTCTGTACTTCTGTCA 900
Db |||||

Search completed: November 10, 2004, 09:24:28
Job time : 608 secs

Qy 980 AGGGG 984
Db 901 CAGGG 905

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 10, 2004, 09:24:31 ; Search time 596 Seconds
(without alignments)
3170.791 Million cell updates/sec

Title: US-10-763-972-2
Perfect score: 1936
Sequence: 1 MSLILLPSRSGSRGAL.....QNKLGHPAGKRCPLNRS 360

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DSV=xlh
-Q=/cgn2_1/USPTO_spool/US10763972/runat_04112004_093313_8778/app_query.fasta_1.519
-DB=N_Geneseq_23Sep04 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10763972@cgn 1.1 470 @runat_04112004_093313_8778 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

	N_Geneseq_23Sep04.*
1:	Geneseqn1980s.*
2:	Geneseqn1990s.*
3:	Geneseqn2000s.*
4:	Geneseqn2001as.*
5:	Geneseqn2001bs.*
6:	Geneseqn2002as.*
7:	Geneseqn2002bs.*
8:	Geneseqn2003as.*
9:	Geneseqn2003bs.*
10:	Geneseqn2003cs.*
11:	Geneseqn2003ds.*
12:	Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1936	100.0	1082	6	ABQ76000 Human GPC
2	1936	100.0	1156	10	ABQ76000 Human GPC
3	1936	100.0	1334	10	ADCS6166 Human GPC
4	1936	100.0	3143	6	AAS17746 Human gen
5	1898	98.0	1063	6	AAS98135 Human DNA
6	1896	97.9	1108	6	ADI16629 Human NOV

7	1896	97.9	1108	12	ADN42283	Adn42283 Human cDN
8	1798	92.9	1002	4	AAS73516	Aas73516 Human G p
9	1792	92.6	1017	10	ADC25999	Adc25999 Human pur
10	1792	92.6	1017	10	AAS61648	Adc61648 Human GPC
11	1785	92.2	1020	6	ABQ79300	Abq79300 Human GPC
12	1785	92.2	1076	6	AAD29667	Ado29667 Human G-p
13	1785	92.2	1076	12	ADO30395	Ado30395 Human GPC
14	1541	79.6	850	6	AAS17747	Aas17747 Human p2y
15	1063.5	54.9	906	12	ADO30397	Ado30397 Mouse GPC
16	921	47.6	510	4	AAS50975	Ahs50975 Human xgp
17	921	47.6	510	6	ABS70208	Abs70208 DNA encod
18	732	37.8	681	4	AAS30774	Aas30774 Human cDN
19	681	35.2	585	4	AAS30783	Aas30783 Human cDN
20	495	25.6	1163	12	ADP49190	Adp49190 Chick p2y
21	460.5	23.8	1086	12	ADO30308	Ado30308 Mouse GPC
22	453	23.4	1163	4	AAD04981	Aad04981 Human pur
23	446.5	23.1	1651	10	ADE40462	Ade40462 Human pyr
24	446	23.0	1177	10	ADS53313	Ads53313 Primary r
25	445.5	23.0	1428	12	ADP21381	Adp21381 Gene p2ry
26	445.5	23.0	1429	2	AAT74321	Aat74321 Human p2y
27	445.5	23.0	1429	10	ACA56470	Aca56470 Human sig
28	445.5	23.0	1429	12	ADI56266	Adi56266 Human pol
29	445.5	23.0	1429	12	ADP49192	Adp49192 Human p2y
30	442.5	22.9	1098	8	ABZ42771	Abz42771 Human pyr
31	442.5	22.9	1098	12	ADO30018	Ado30018 Human GPC
32	442	22.8	1134	12	ADO30017	Ado30017 Human GPC
33	442	22.8	2025	4	AAD04980	Aad04980 Human pur
34	442	22.8	2025	6	ABZ35611	Abz35611 Human gen
35	442	22.8	2025	8	ABZ42713	Abz42713 Human pur
36	442	22.8	2025	10	ACA56884	Aca56884 Human sig
37	442	22.8	2025	12	ADI56680	Adi56680 Human pol
38	442	22.8	2025	12	ADP49188	Adp49188 Human p2y
39	440	22.7	2118	6	ABZ35320	Abz35320 Human gen
40	436	22.5	1922	10	ABT42366	Abt42366 Toxicity
41	435	22.5	2547	12	ADP21380	Adp21380 Gene p2ry
42	420.5	21.7	984	2	AAT18367	Aat18367 Mouse pan
43	420.5	21.7	987	12	ADO30309	Ado30309 Mouse GPC
44	420.5	21.7	1881	10	ADC37338	Adc37338 Nuclear f
45	416	21.5	1842	2	AAQ86134	Aaq86134 Human p2o

ALIGNMENTS

RESULT 1	ABQ76000	ABQ76000 standard; DNA; 1082 BP.
ID	ABQ76000	
XX	XX	
AC	ABQ76000;	
XX	XX	
DT	17-OCT-2002	(first entry)
DE	Human GPCR designated PFI-020 encoding sequence.	
XX	XX	
KW	Human; GPCR; G-protein coupled receptor; antidepressant; neuroleptic;	
KW	gene therapy; therapeutic; mood; depression; arousal; eating; sleeping;	
KW	disorder; PFI-020; gene; ds.	
OS	Homo sapiens.	
XX	XX	
FH	Key	Location/Qualifiers
FT	CDS	1..1080
FT	FT	/tag= a
FT	FT	/partial
FT	FT	/product= "GPCR designated PFI-020"
FT	FT	/note= "no stop codon given"
XX	XX	
PN	EP1215214-A1.	
XX	XX	
PD	19-JUN-2002.	
XX	XX	
PF	04-DEC-2001; 2001EP-00310137.	
XX	XX	
PR	18-DEC-2000; 2000GB-00030855.	

DR WPI; 2003-364675/35.
 DR P-PSDB; ADB99264.
 XX
 PT New human gene P2Y2L1 and encoded G protein-coupled receptor, useful for
 PT treatment and diagnosis of receptor-associated diseases and for drug
 PT screening.
 XX
 PS Disclosure; Page 4; 5pp; German.
 XX
 CC This invention describes the human P2Y2L1 gene and its 5' - and 3' -
 CC untranslated regions, located on chromosome 3 which is a novel G protein-
 CC coupled receptor (GPCR). The protein encoded by P2Y2L1 is expressed in
 CC blood cells, testis and embryonal kidney cells and contains potential
 CC sites for phosphorylation by protein kinase C and casein kinase II. It is
 CC a member of the Class A rhodopsin-like sub-family of G protein-coupled
 CC receptors and it probably a nucleoside/nucleotide receptor that mediates
 CC action of nucleosides/nucleotides or their sugar derivatives. P2Y2L1 and
 CC antibodies directed against the encoded protein are useful in diagnosis
 CC and treatment (including gene therapy) of diseases associated with
 CC abnormal levels of P2Y2L1 expression, in screening assays for modulators,
 CC potential therapeutic agents, and to produce transgenic animals, e.g. for
 CC identifying diseases associated with abnormal expression of P2Y2L1. This
 CC sequence represents the P2Y2L1 gene described in the disclosure of the
 CC invention.
 XX
 SQ Sequence 1156 BP; 243 A; 335 C; 295 G; 283 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,2e-176 Length: 1156
 Score: 1936.00 Matches: 360
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-10-763-972-2 (1-360) x ADB99267 (1-1156)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
 DB 21 ATCTGTGCAATTTTCTCTCCAGGGAAGCAGAGGGGAGCGTGTGGAGCTGTG 80
 QY 21 LeuLeuGluAlaSerArgAspMetGluLeuValAspMetAsnThrSerGlnGluGln 40
 DB 81 CTCTGTGAGGAGCGCTCCCGGAGCATGGAGAGGTGGACATGAATACATCAGAGAACAA 140
 QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
 DB 141 GGTCTGTGCGAGTCTTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 200
 QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheSerTrpGlyGlnThr 80
 DB 201 ATCTTTATCTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 260
 QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
 DB 261 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCAGCTCTTTAT 320
 QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpPheGlyGlu 120
 DB 321 GTGCTATTGCCCTTCTCATCATCACTACTACTAGTACAGAGTGGCCCTTCGGGGAG 380
 QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
 DB 381 CTGCTCTGCAAGTGGTGGCACTTCTCTGTATATCAACCTTTACGGCAGCATCTCGCTG 440
 QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 DB 441 CTGACCTGCATCTCTGTGCCACAGTCTCTAGGTGTGGCCACCCACTGTGTCTGCTGCC 500
 QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeu 180
 DB 501 TACCGGACCGGAGCATGTGCTGGTGGCAGCAGCACCCACCTGGGCGCTGTGGTCTC 560

QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
 DB 561 CAGCTGTGTGCCACACTGGGCTTCTCCACAGGACTACATCAATGATGCCAGATGATCTGG 620
 QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
 DB 621 TATGACATGACCAAGAGATTTTGTATCGGCTTTTGTCTACGGCATAGTCTTGACA 680
 QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 DB 681 TTGTCTGGCTTCTTCTCCCTTGGTCATTTTGTGTGTCTATTCATGATGGTCAGAG 740
 QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
 DB 741 CCTGATCAGCCAGGAGGAACTCATGAGGACAGGCAACACAGCCCGAGCGAGTCCAT 800
 QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
 DB 801 CCGGACCATCTACTGTGTGTGGCTCTTCACTCTGTGTGTGTGTGTGTGTGTGTGTGT 860
 QY 281 SerLeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 DB 861 TCGTCTCTCTTACTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 920
 QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
 DB 921 AGCCAGTGTGGCTTACAGATATGAGAGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGT 980
 QY 321 SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
 DB 981 AGTCTGTACTTCTTTCAGGGGGGCAAAATAGATGAGCTCTCTCCAGAAACTGAGG 1040
 QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
 DB 1041 CAGAACAGTTGGGTGAGCATCCAGTGGGAGGAAGATGCCCGAGGTTTGAACAGATCT 1100

RESULT 3
 ADC86166
 ID ADC86166 standard; DNA; 1334 BP.
 XX
 AC ADC86166;
 XX
 DT 01-JAN-2004 (first entry)
 DE Human GPCR gene SEQ ID NO:619.
 XX
 KW ds; gene; human; GPCR;
 XX
 KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1270724-A2.
 XX
 PD 02-JAN-2003.
 XX
 PF 18-JUN-2002; 2002EP-00013517.
 XX
 PR 18-JUN-2001; 2001JP-00246789.
 XX
 PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 XX
 PI Suwa M, Asai K, Akiyama Y, Aburatani H;
 XX
 DR WPI; 2003-315783/31.
 DR P-PSDB; ADC86167.
 XX
 PT New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT Guanosine triphosphate-binding protein coupled receptor.
 XX
 PS Claim 1; SEQ ID NO 619; 28pp; English.

CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The
 CC polynucleotide sequences shown in ADC8548-ADC87616 encode GPCR's of the
 CC invention.

XX
 SQ Sequence 1334 BP; 283 A; 368 C; 348 G; 335 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,48e-176 Length: 1334
 Score: 1936.00 Matches: 360
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

us-10-763-972-2 (1-360) x ADC86166 (1-1334)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
 DB 117 ATGCTGTCCATTTCCTTCCTTCAGGGGAAGCAGAGCGGAGCGCTGCTGGAGCTCTG 176
 QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
 DB 177 CTCCTGAGGGAGCCTCCGGGACATGGAGAGGTGGACATGAATACATCACAGGAACAA 236
 QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
 DB 237 GGTCTCTGCCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 296
 QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheSerIleGlnThr 80
 DB 297 ATCTTTATCTAGGCTGGCACTAAATGGACACTGCTTGTGGCACTCTCTGGGCGCAAC 356
 QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
 DB 357 AAGCGCTGGAGCTGTGCCACCATCTCTGTGTGAACCTGATGTGTGGCGCACCTGCTTTAT 416
 QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspAspArgTrpPheGlyGlu 120
 DB 417 GTGCTATTGCCCTTCCTATCATCATCCTACTACTAGATGACAGGTGGCCCTTCGGGGAG 476
 QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
 DB 477 CTGCTCTGCAAGCTGTGGCACTCTCTCTATATCAACCTTTACGGCAGCATCTCTGTG 536
 QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 DB 537 CTGACCTGCACTCTGTGCAACAGTTCCTAGGTGTGTGCCACCACTGTTCGCTGCCCC 596
 QY 161 TyrArgThrArgHisAlaThrLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
 DB 597 TACCGGACCCGAGGCATGCTGCTGGGCACCGACCACTGCTGGCCCTGTGTGTCTCTC 656
 QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
 DB 657 CAGCTGTGCCCACTGTGCCCTTCTCCACACGGACTATCAATGCGCAGATGATCTGTG 716
 QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
 DB 717 TATGACATGACCGAGCAAGAGATTTTGATCGGCTTTTGGCTACGGCATAGTCTTGACA 776
 QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 DB 777 TTGCTGTGGCTTCTTTCTCCCTCTGTGTGTCATTTTGTGTGTCATTCTGATGTCAGGAG 836
 QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
 DB 837 CCTGATCAACCCAGAGGAGAACCTCATGAGGACAGGACACACAGCCCGGAGCCAGTCCAT 896

QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
 DB 897 CGGACCACTCTACTGGTGTGGCTTTCACCTCTGTTTGTGGCTTCCATATCAC 956
 QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 DB 957 TCGCTCTCTTACCTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGC 1016
 QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
 DB 1017 AGCCAGTGTGCCCTACAGATATGGAGGCTCTGCTGAGTGTGAGCAGCTCCCAACCC 1076
 QY 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
 DB 1077 AGTCTGTGACTTCTTCAAGGGGGCAAAATAGAGTCAGGCTCTCCAGAACTGAGG 1136
 QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
 DB 1137 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCCGAGGTGAACAGATCT 1196

RESULT 4

AAS17746
 ID AAS17746 standard; DNA; 3143 BP.

XX AAS17746;
 AC AAS17746;

DT 26-FEB-2002 (first entry)

XX Human genomic clone for P2Y-like G protein-coupled receptor.

XX Human; ds; P2Y-like G protein-coupled receptor; GPCR; COPD;
 KW Chronic obstructive pulmonary disease; nervous system disease;
 KW Parkinson's disease; multiple sclerosis; dementia; stroke;
 KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;
 KW bacterial infection; fungal infection; protozoan infection;
 KW viral infection; pain; cancer; anorexia; bulimia; asthma;
 KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;
 KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;
 KW psychotic disorder; neurological disorder; anxiety; schizophrenia;
 KW manic depression; delirium; severe mental retardation; dyskinesia.

XX Homo sapiens.

OS
 FH Key Location/Qualifiers
 CDS 520..2498

FT /*tag= a
 FT /product= "P2Y-like GPCR"

XX WO200185764-A2.

XX 15-NOV-2001.

XX 09-MAY-2001; 2001WO-EP005244.

XX 11-MAY-2000; 2000US-0203582P.

XX 21-FEB-2001; 2001US-0269857P.

XX (FARB) BAYER AG.

XX Ramakrishnan S;

XX WPI; 2002-075242/10.

XX P-PSDB; AAU11251.

XX New polynucleotides for producing P2Y-like G protein-coupled receptors
 PT (GPCR) that are used for screening inhibitors or regulators of human P2Y-
 PT like GPCR, especially useful for treating pain, cancer or neurological
 PT disorders.

XX Disclosure; Fig 1; 114pp; English.

XX The invention relates to an isolated polynucleotide encoding a P2Y-like G
 CC protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or

CC allele, a host cell containing an expression vector comprising the
CC polynucleotide and screening for agents that regulate the GPCR activity.
CC The polynucleotide is useful for producing P2Y-like GPCR polypeptide,
CC which may be employed for screening agents that inhibit or regulate human
CC P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is
CC useful for treating or ameliorating P2Y-like GPCR disorders, particularly
CC COPD (chronic obstructive pulmonary disease), peripheral or central
CC nervous system disease (e.g. Parkinson's disease, multiple sclerosis,
CC dementia, stroke, Alzheimer's disease and many other diseases and
CC disorders listed in the specification), benign prostatic hyperplasia or
CC urinary incontinence. A pharmaceutical composition containing the
CC modulators and/or regulators of P2Y-like GPCR is useful for modulating
CC the activity of a P2Y-like GPCR. In particular, these are useful for
CC treating, preventing or ameliorating infections (e.g. bacterial, fungal,
CC protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma,
CC acute heart failure, hypotension, hypertension, osteoporosis, diabetes,
CC angina pectoris, myocardial infarction, ulcers, inflammation, allergies,
CC psychotic or neurological disorders (e.g. anxiety, schizophrenia, manic
CC depression, delirium, severe mental retardation or dyskinesias). The
CC present sequence is a genomic clone encoding the P2Y-like GPCR of the
CC invention

SQ Sequence 3143 BP; 818 A; 771 C; 726 G; 828 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,07e-176 Length: 3143
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-763-972-2 (1-360) x AAS17746 (1-3143)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
DB 436 ATGCTGCTCCATTGCTTCTCCAGGGAGACAGAGAGCGGCGCTGCTGAGCTCTG 495
QY 21 LeuLeuGluGlyAlaSerArgMetGluLysValAspMetAsnThrSerGlnGlu 40
DB 496 CTCTGGAGGGAGCCCTCCGGGACATGGAGAGGTGGACATGATACATCACAGAACAA 555
QY 41 GlyLeuSerGlnPheSerGluLysValGlnValTyrLeuSerLeuAlaTyrSerile 60
DB 556 GGTCTCTGCAGTCTTCAGAGAGTACAGCAGTCTACTCTCTCCGCGCTACAGTATC 615
QY 61 IlePheIleLeuGlyLeuProLeuAenGlyThrValLeuTrpPheSerTrpGlyGln 80
DB 616 ATCTTTATCTAGGGCTGCCACTTAATGGCACTGTCTGTGGCACTCTCTGGGCGCAACC 675
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuTyr 100
DB 676 AGCGGTGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGTTTAT 735
QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120
DB 736 GTGCTATTGCCCTTCTCATCATCACTACTACATAGATGACAGTGGGCCCTCGGGAG 795
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeu 140
DB 796 CTGCTCTGCAAGCTGTGGCTCTCTCTATATCAACCTTTACGGCAGCATCTCTGCTG 855
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
DB 856 CTGACCTGATCTCTGTGCACAGTCTTAGGTGTGTGCCACCCACTGCTGTGTGCTGCC 915
QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
DB 916 TACCGGACCGCAGGATGCTGCTGGGACAGTCCAGTGTGTGCCACCCACTGCTGTGTGCTCC 975
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
DB 976 CAGTGTGCTCCACACTGGCTCTCTCCACAGGATACATCAATATGGCAGATGATCTGG 1035

QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
DB 1036 TATGACATGACACAGCCAGAGAAATTTGATCGGCTTTTTCCTACGCATAGTTCGACA 1095
QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
DB 1096 TTGCTGCTGCTTCTTCTCCCTCTTGGTGCATTTTGGTGTGCTATTCACTGATGTCAGAG 1155
QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
DB 1156 CCGTATCAAGCCAGGAGAACCTCATGAGGACAGGACACAGCCCGCCAGGTCAT 1215
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
DB 1216 CCGGACCATCTACTGCTGTGTGGCTCTTCACTCTGCTTTTGTGGCTTCCATATCAC 1275
QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
DB 1276 TCGCTCTCTTACCTCACCATCTGCTTCTTCTCAGGACTGCCAGCTCTTGATGGC 1335
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
DB 1336 AGCCAGTGTGGCTTACAGATATGGAGGCTCTGTGTGAGTGTGAGCAGCTGCCTCAACC 1395
QY 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
DB 1396 AGTCTCTGCTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCTCCAGAACTCAGG 1455
QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
DB 1456 CAGAACAGTTGGTGTGAGCATCCAGCTGGAGAGAGAGATGCCAGGCTTGACAGATCT 1515

RESULT 5

QY AAS98135 ID AAS98135 standard; DNA; 1063 BP.
DB XX AC AAS98135;
XX XX 12-MAR-2002 (first entry)
DE Human DNA for potential G protein-coupled receptor #92.
KW Human; G protein-coupled receptor; GPCR; Gs; GAL4; galanin receptor;
KW Alzheimer's disease; amyotrophic lateral sclerosis; asthma;
KW atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;
KW chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;
KW depression; epilepsy; macular degeneration; lymphoma; melanoma;
KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;
KW psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;
KW tuberculosis; cognition disorder; memory disorder; anorexia;
KW hormonal release disorder; cardiovascular activity disorder;
KW pain perception disorder; obesity; diabetes; obesity; diabetes;
KW hyperlipidaemia; stroke; gene therapy.
OS Homo sapiens.
FN WO200185791-A1.
PD 15-NOV-2001.
XX 11-MAY-2001; 2001WO-US015332.
XX 11-MAY-2000; 2000US-0203217P.
XX 18-MAY-2000; 2000US-0205945P.
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
PI Brown JP, Miller M, Burmer G, Fabre-Suver C, Pritchard D;
XX WPI; 2002-066595/09.
XX Novel G protein-coupled receptor polypeptides including galanin receptor

PT polypeptides useful for identifying modulators that are useful for
PT treating Alzheimer's disease, psoriasis, melanoma, multiple sclerosis,
XX stroke.

PS Claim 2; Page 98; 144pp; English.

XX The invention relates to an isolated polypeptide encoded by a nucleic
CC acid molecule that is at least 80% identical to the G protein-coupled
CC (GPCR) polynucleotides included in the specification. Also included are
CC probes based on the GPCR sequences (including antisense probes), a host
CC cell comprising an expression vector comprising the GPCR sequence,
CC antibodies raised against the polypeptides, and methods of identifying
CC modulators of the polypeptides. The polypeptides are useful for
CC identifying modulator compounds which function as modulators, activators,
CC repressors, agonists or antagonists of the novel GPCR polypeptides
CC including the GAL4 polypeptide. The antibodies and nucleic acid probes as
CC described above can be used to detect the presence of the polypeptides
CC and nucleic acids and are used to diagnose a variety of diseases or
CC disorders in which GPCRs are involved e.g., Alzheimer's disease,
CC amyotrophic lateral sclerosis, asthma, atherosclerosis, basal cell
CC carcinoma, breast carcinoma, cardiomyopathy, chondrosarcoma, chronic
CC obstructive pulmonary disease, Crohn's disease, depression, epilepsy,
CC macular degeneration, lymphoma, melanoma, multiple sclerosis,
CC osteoarthritis, osteoporosis, Parkinson's disease, psoriasis, rheumatoid
CC arthritis, schizophrenia, ulcerative colitis, tuberculosis and many other
CC diseases listed in the specification. The probes and antibodies are also
CC useful for diagnosing cognition and memory disorders, anorexia, hormonal
CC release disorders, cardiovascular activity disorders, pain perception
CC disorders, obesity, diabetes, Alzheimer's disease. Preferably, compounds
CC that decrease or increase the expression of galanin receptor (GAL4) can
CC be used to treat obesity, diabetes, hyperlipidaemia and stroke. The GPCR
CC nucleic acid is useful for treating the above mentioned disorders by gene
CC therapy techniques. The present sequence is a novel GPCR polynucleotide
CC of the invention

XX SQ Sequence 1063 BP; 216 A; 315 C; 272 G; 260 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.92e-173 Length: 1063
Score: 1898.00 Matches: 354
Percent Similarity: 99.72% Conservatives: 0
Best Local Similarity: 99.72% Mismatches: 0
Query Match: 98.04% Indels: 1
DB: 6 Gaps: 0

US-10-763-972-2 (1-360) x AAS98135 (1-1063)

Qy 7 ProSerHrglySerArgSerGlySerArgArgGlyAlaLeuLeuGlyAlaSer 26
Db 1 CCTTCCAGGGGAAGCAGAGCGGAGCCGCTGCTGGAGCTCTGCTGGAGGAGCCTCC 60
Qy 27 ArgAspMetGluLysValAspMetAsnThrSerGlnGlnGlnGlyLeuGlyLeuGlyLeu 46
Db 61 CGGACATGAGAGGAGTGACATACATACATACAGGACAGGCTCTGCGCAGTTCTCA 120
Qy 47 GluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheLeuGlyLeu 66
Db 121 GAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACATATCATATCTTTATCTTACCTAGGGCTG 180
Qy 67 ProLeuAsnGlnThrValLeuTyrPheSerTrpGlyGlnThrLysArgTrpSerCysAla 86
Db 181 CCATAAATGGCACTGCTTTGGCACTCTCTGGGCGCAACCAAGCGCTGGAGCTGTGCC 240
Qy 87 ThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuPropheLeu 106
Db 241 ACCACCTATCTGGTGAACCTGATGGTGGCGCACTGCTTTATGCTATATGCTTCTCTC 300
Qy 107 IleIleThrTyrSerLeuAspAspArgTrpPheGlyGlnLeuLeuCysLysLeuVal 126
Db 301 ATATCATCTACTACTAGATGACAGTGCGCTCTGGGAGCTGCTCTGCAAGCTGGTG 360
Qy 127 HisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerVal 146

Db 361 CACTTCTGTTCTATATCAACCTTTAGCGGAGCATCTCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 147 HisGlnPheLeuGlyValCysHisProLeuLeuCysSerLeuProTyrArgThrArgHis 166
Db 421 CACCAGTTCTTAGGTGTGTCACCCACTGTGTGTGCTGCTTACCGGACCGGAGCAT 480
Qy 167 AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuLeuGlnLeuProThrLeu 186
Db 481 GCCTGGCTGGGACACAGCACCATCTGGGCTGTGTGTCTCTCCAGCTGTGCTCCCACTG 540
Qy 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206
Db 541 GCCTTCCTCCACACGGACTACATCAATGGCCAGATGATCTGGTATGATGATGACGAGCAA 600
Qy 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSer 226
Db 601 GAGAATTTTGATCGGCTTTTGGCTACGGCATAGTTCTGACATTGTCTGGCTTTCTTTTC 660
Qy 227 LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGly 246
Db 661 CTCTTGGTCAATTTGGTGTGCTATTCTACTGATGGTCAGAGGCTGATCAAGCCAGGGA 720
Qy 247 GluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGly 266
Db 721 GAACCTCATGAGGACAGGCAACACAGCCGAGGCTCCATCCGCGACCATCTTACTTGGT 780
Qy 267 ValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuProHis 286
Db 781 GTGTGGCTCTTCACTCTGTTGTGGCTTCCATATCATCTGCTCTCTTACTTCTACCTC 840
Qy 287 HisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuG 306
Db 841 CATCTGCTTTCTGCTTCTCAGGACTCCAGCTCTTGTATGCGACGCCAGTGTGCGCTACA 900
Qy 306 AspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPh 326
Db 901 AGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAACCCAGTCTCTGCTGCTTTCTTT 960
Qy 326 eLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyG 346
Db 961 CAAGGGGGCAAAATAGATGATGAGCTCTCCAGAACTGAGGAGGAGCAAGTTGGGTGA 1020
Qy 346 uHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
Db 1021 GCATCCAGCTGGGAGGAGATGCCAGGCTTGAACAGATCT 1063

RESULT 6
ADI16629 standard; cDNA; 1108 BP.
XX AC ADI16629;
XX DT 15-APR-2004 (first entry)
XX DE Human NOVX cDNA to treat human pathological conditions SeqID165.
XX gene; ss; human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; stroke; muscular dystrophy; epilepsy;
KW cytosolic; cardiac; antiinflammatory; immunosuppressive; anorectic;
KW haemostatic; anti-HIV; antidiabetic; antiarteriosclerotic; hepatocytic;
KW antiaslomatic; nephrotropic; antithrombotic; hepatocytic;
KW neuroprotective; nootropic; antibacterial; virucide; antiparasitic;
KW relaxant; anticonvulsant; neurogenesis; wound healing; angiogenesis;
KW chromosome mapping; tissue typing; pharmacogenomic;
KW single nucleotide polymorphism; SNP.
XX Homo sapiens.
XX CS
XX Key variation
XX Key Location/Qualifiers
XX replace (378,A)
XX FT

PA (RAST/) RASTELLI L.
 PA (MILL/) MILLER C E.
 PA (GERL/) GERLACH V.
 PA (TAUP/) TAUPIER R J.
 PA (GUSE/) GUSEV V Y.
 PA (COLM/) COLMAN S D.
 PA (WOLE/) WOLENC A R.
 PA (PENA/) PENNA C E A.
 PA (FURT/) FURTA K.
 PA (GROS/) GROSSE W M.
 PA (ALSO/) ALSOBROOK J P.
 PA (LEPL/) LEPLEY D M.
 PA (RIEG/) RIEGER D K.
 PA (BURG/) BURGESS C E.
 XX
 PI Tchernev VT, Spytek KA, Zernhusen BD, Patturajan M, Shimkets RA;
 PI Li L, Ganggoli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;
 PI Gerlach V, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;
 PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
 XX
 DR WPI; 2004-180039/17.
 DR P-PSDB; ADN42284.
 XX
 PT Isolated NOVX polypeptides and polynucleotides, useful for preventing
 PT diagnosing and/or treating cancer, diabetes, cardiomyopathy and
 PT atherosclerosis.
 XX
 PS Claim 9; SEQ ID NO 165; 1309pp; English.
 XX
 CC The invention relates isolated 162 NOVX polypeptides (NOVX-NOV99,
 CC including splice variants) and the nucleic acids (NA) that encode them.
 CC Also included are the mature NOVX proteins (NA) and their encoding
 CC polynucleotides), a vector comprising NOVX NA, a cell comprising the
 CC vector, an antibody that binds immunospecifically to NOVX, determining
 CC the presence or amount of NOVX in a sample, determining the presence or
 CC amount of NOVX NA in a sample, identifying an agent that binds to NOVX,
 CC modulating the activity of NOVX, treating or preventing a NOVX-associated
 CC disorder, determining the presence of or predisposition to a disease
 CC associated with altered levels of NOVX and treating a pathological state
 CC in a mammal comprising administering a polypeptide which is at least 95%
 CC identical to NOVX (or fragment). NOVX and NA may be used in the
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate expression and activity of NOVX (e.g. cancer, diabetes,
 CC cardiomyopathy and/or atherosclerosis). The anti-NOVX antibodies and
 CC antagonists may also be used to down regulate expression and activity of
 CC NOVX. The anti-NOVX antibodies may also be used as diagnostic agents for
 CC detecting the presence of NOVX in samples (e.g. by enzyme linked
 CC immunosorbent assay (ELISA). The agents and methods may be used in this
 CC way to prevent, diagnose and treat cancer, diabetes, cardiomyopathy
 CC and/or atherosclerosis. The present sequence encodes a NOVX protein.
 XX
 SQ Sequence 1108 BP; 225 A; 321 C; 285 G; 277 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 8.14e-173 Length: 1108
 Score: 1896.00 Matches: 359
 Percent Similarity: 99.45% Conservative: 0
 Best Local Similarity: 99.45% Mismatches: 1
 Query Match: 97.93% Indels: 2
 DB: 12 Gaps: 0

US-10-763-972-2 (1-360) x ADN42283 (1-1108)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
 DB 12 ATGCTGTCATTTTGTCTTCCTCCAGGGGAAGCAGAGCGGAGCGCTGCTGGAGCTCTG 71
 QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
 DB 72 CTCCTGGAGGAGCGCTCCCGGACATGAGAGGTGGACATGATATACATCACAGGAACAA 131
 QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60

RESULT 8

AAH73516
 ID AAH73516 standard; DNA; 1002 BP.

XX
 AC AAH73516;
 XX

DT 25-SEP-2001 (first entry)

DB 132 GGTCTGCCAGTTCTCAGAGAAAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 191
 QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
 DB 192 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGGCAAAACC 251
 QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
 DB 252 AAGGCTGGAGCTGTGCCACCACCTATCTGTGTGAACCTGTATGGTGGCGACGCTCTTAT 311
 QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120
 DB 312 GTGCTATTGCCCTTCTCTCATCATCACTACTCATAGATGACAGGTGGCCCTTCGGGGAG 371
 QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
 DB 372 CTGCTCTGCAAGCTGGTGCACTTCTGTCTATATCAACCTTTACGGCAGCATCCTGCTG 431
 QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 DB 432 CTGACCTGCATCTCTGTGCACCACTTCTTAGTGTGTGCCACCACCACTGTGTCTGCTGCC 491
 QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
 DB 492 TACCGGACCCGACGACATGCTGCTGGCACCACGACCACTGGGCGCTGCTGCTGCTC 551
 QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
 DB 552 CAGCTGCTGCCACACTGGCCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 611
 QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
 DB 612 TATGACATGACAGCAAGAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGACA 671
 QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 DB 672 TTGCTGCGCTTCTTCTCCCTCTGCTGCTATTTGGTGTA-TTCACCTGATGGTCAGGAG 730
 QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
 DB 731 CCTGATCAAGCAGAGGAGAACCTCATGAGGACAGGCAACACAGCCGAGCCAGTCCAT 790
 QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
 DB 791 CCGGACCATCTCTACTGCTGTGTGGCTTTCACCTCTCTGTTTGTGCCCTTCCATATCAC 850
 QY 281 SerLeuLeuLeuProHisHisIleLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 DB 851 TGGCTCTCTTACCTTCACCACTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGTATGGC 910
 QY 301 Ser-GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPr 320
 DB 911 AGCCCAAGTGTGGCTTACAAGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCCTCAACC 970
 QY 320 OSerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuAr 340
 DB 971 CAGTCTCTTACTTCTTCAAGGGGGGCAAAAATAGACTAGGCTCTCCCAAAACTCAG 1030
 QY 340 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSe 360
 DB 1031 GCAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCCGAGGTTGAACAGATC 1090
 QY 360 r 360
 DB 1091 T 1091.

XX DE Human G protein-coupled receptor GPRv71 coding sequence.

XX KW Human; guanosine triphosphate binding protein-coupled receptor;

KW G protein-coupled receptor; GPRv8; GPRv12; GPRv16; GPRv21; GPRv40;

KW GPRv47; GPRv51; GPRv72; cancer; liver cirrhosis;

KW Alzheimer's disease; cytostatic; hepatotropic; neotropic;

KW neuroprotective; gene therapy; peptide therapy; ds.

OS Homo sapiens.

XX WO200148188-A1.

XX 05-JUL-2001.

XX 28-DEC-2000; 2000WO-JP009408.

XX 28-DEC-1999; 99JP-00375152.

XX 31-MAR-2000; 2000JP-00101339.

XX (HELI-) HELIX RES INST.

XX Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;

XX Sugiyama T, Kishimoto T, Kanzaki K, Yasuda S, Inoue Y;

XX WPI; 2001-425662/45.

XX P-ESDB; AAG64125.

XX New DNA encoding guanosine triphosphate binding protein coupled receptors

XX and their expression products for screening potential anticancer and

XX neotropic drugs and in diagnosis of these diseases.

XX Example 1; Page 142-143; 170pp; Japanese.

XX The invention relates to nine human guanosine triphosphate binding

XX protein (G protein)-coupled receptors designated GPRv8, GPRv12, GPRv16,

XX GPRv21, GPRv40, GPRv47, GPRv51, GPRv71 and GPRv72, and to the genes

XX encoding them. These genes and proteins and antibodies against the

XX protein are useful in the treatment, prevention, diagnosis and

XX investigation of diseases associated with G protein-coupled receptors,

XX including cancer, cirrhosis of the liver and Alzheimer's disease. The

XX present sequence is a G protein-coupled receptor coding sequence of the

XX invention

XX SQ Sequence 1002 BP; 207 A; 293 C; 249 G; 253 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.98e-163	Length:	1002
Score:	1798.00	Matches:	331
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	92.87%	Indels:	0
DB:	4	Gaps:	0

US-10-763-972-2 (1-360) x AAH73516 (1-1002)

QY 29 MetGluValAspMetAsnThrSerGlnGluGlnGlyLeuCysGlnPheSerGluLys 48

DB 1 ATGGAGAGGTGGACATGATATACATACAGAGACAGGCTCTGCCAGTCTCAGAGAG 60

QY 49 TyrIysGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeu 68

DB 61 TACAAGCAAGTCTACCTCTCCCTGGCTACAGATATCATCTTTATCTAGGCTGCCACTA 120

QY 69 AsnGlyThrValLeuTyrHisSerTrpGlyGlnThrIleArgTrpSerCysAlaThrThr 88

DB 121 AATGGCACTGCTTGTCGCACTCTCTGGGGCAACCAAGCGCTGGAGCTGGCCACC 180

QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeuIleIle 108

DB 181 TATCTGGTGAACCTGATGGTGGCGACCTGCTTTATGTGCTATTGGCCCTTCCTCATCATC 240

QY 109 ThrTyrSerLeuAspAspArgTrpProPheGlyGluLeuLeuCysIysLeuValHisPhe 128

DB 241 ACCTACTCCTAGATGACAGGTGGCCCTTCGGGAGCTGCTCTCAAGCTGGTGACATTC 300

QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148

DB 301 CTGTTCTATATACACCTTTAGCGGAGCATCTGCTGCTGACCTCATCTCTGTGCACAG 360

QY 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTrp 168

DB 361 TTCCTAGGTGTGGCCACCCACTGCTGTTGGCTGCCCTACCGGACCGGAGCATGCTGG 420

QY 169 LeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188

DB 421 CTGGGCACACAGCACCTGGGCCCTGGTGGTCTCCAGCTGCTGCCACACTGGCCTTC 480

QY 189 SerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGluAsn 208

DB 481 TCCACACGAGCTACATCATGGCCAGATGATCTGGTATGACATGACACGACAGAGAT 540

QY 209 PheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 228

DB 541 TTGATCGGCTTTTGGCTACGGCATAGTCTGACATGTCTGGCTTCTTTCCTCTCTT 600

QY 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248

DB 601 GGTCACTTTGGTGTGCTATTCACTGATGTCAGGAGCTGATCAAGCCAGAGGAGACCT 660

QY 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268

DB 661 CATGAGGACAGGACACAGCCGAGCCAGGTCCATCCGGACCATCTTACTGTGTGTGG 720

QY 269 ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisIleLeu 288

DB 721 CCTCTCACCTCTGTTTGTGGCTTCCATATCACTCGCTCTCTTACCTCACCATCTG 780

QY 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308

DB 781 CTTTCTGTTTCTCAGGACTGCGAGCTTTGATGGCAGCCAGTGTGGCTACAGATATG 840

QY 309 GluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPheGlyGly 328

DB 841 GAGGCTCTGGTGTGTGTGAGCAGCTGCTCAAGCCAGTCTCTGTACTTCTTCTTCAAGGG 900

QY 329 GlyIysAsnArgValArgLeuLeuGlnIysLeuArgGlnAsnLeuLeuGlyGluHisPro 348

DB 901 GCATAAATAGATCAGGCTCTCTCAGAACTGAGGAGAGACAGATTGGGTGACATCCA 960

QY 349 AlaGlyArgLysArgCysProGlyLeuAsnArgSer 360

DB 961 GCTGGAGAGAGAGATGCCAGGCTTGAACAGATCT 996

RESULT 9

ADC25999

XX ADC25999 standard; DNA; 1017 BP.

XX AC ADC25999;

XX 18-DEC-2003 (first entry)

XX Human purinergic receptor P2Y-related GPCRv6 DNA.

XX virucide; fungicide; antibacterial; cytostatic; analgesic; antidiabetic;

XX anorectic; cardiant; hypotensive; osteopathic; antiangiinal;

XX antiarteriosclerotic; cerebroprotective; anti-ulcer; antiallergic;

XX neotropic; neuroprotective; antiparkinsonian; G-protein coupled receptor;

XX GPCR; viral; fungal; bacterial infection; immune-related disorder;

XX cancer; pain; diabetes; obesity; anorexia; acute heart failure;

XX hypertension; osteoporosis; angina pectoris; atherosclerosis; stroke;

XX ulcer; allergy; psychotic neurological disorder; schizophrenia; dementia;

XX degenerative disease; Parkinson's; Alzheimer's; dyskinnesia; Huntington's;

XX human; GPCRv6; purinergic receptor P2Y; ds; gene.

XX Homo sapiens.

```

XX Key
FH CDS
FT Location/Qualifiers
FT 1..1017
FT /tag= a
FT /product= "Human purinergic receptor P2Y-related GPCRx6
FT protein - both "original" and "alternative" versions"
FT /transl_except= (pos:184..186, aa:Pro)
FT /transl_except= (pos:727..729, aa:Xaa)
FT /transl_except= (pos:808..810, aa:Pro)
FT /note= "Xaa = Unknown; translation exceptions at 184 and
FT 727 are present within "original" protein CDS whilst that
FT at 808 is present within the "alternative" protein CDS"
FT
XX US2003089080-A1.
XX
XX 08-MAY-2003.
XX
XX 21-JUN-2001; 2001US-00895453.
XX
XX 20-JUN-2000; 2000US-0212908P.
XX 05-DEC-2000; 2000EP-00870289.
XX
XX (COMM/) COMMUNI D.
XX (LANV/) LANNOY V.
XX (GOVA/) GOVAERTS C.
XX (PARM/) PARMENTIER M.
XX (DETH/) DETHEUX M.
XX
XX Communi D, Lannoy V, Govaerts C, Parmentier M, Detheux M;
XX WPI; 2003-657993/62.
XX P-PSDB; ADC26000, ADC26009.
XX
XX New human G-protein coupled receptor, useful for treating receptor-
XX mediated disorders, e.g. infections, cancer, pain, diabetes, obesity,
XX acute heart failure, osteoporosis, stroke, ulcer, allergy, or
XX neurological disorders.
XX
XX Example 3; Page 15-16; 24pp; English.
XX
XX The invention relates to a novel G-protein coupled receptor (GPCR). The
XX receptor, polynucleotide, agonist, reverse agonist and antagonist of the
XX invention may be useful for treating receptor-mediated disorders
XX including viral, fungal or bacterial infections, immune-related disorders
XX such as cancer, pain, diabetes, obesity, anorexia, acute heart failure,
XX hypertension, osteoporosis, angina pectoris, atherosclerosis, stroke,
XX ulcer and allergy, as well as psychotic and neurological disorders such
XX as schizophrenia and dementia, degenerative diseases such as Parkinson's
XX disease and Alzheimer's disease and dyskinesias such as Huntington's
XX disease. The current sequence is that of the human purinergic receptor
XX P2Y-related GPCRx6 DNA of the invention.
XX
XX SQ Sequence 1017 BP; 200 A; 306 C; 252 G; 259 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 7.68e-163 Length: 1017
XX Score: 1792.00 Matches: 338
XX Percent Similarity: 99.71% Conservative: 0
XX Best Local Similarity: 99.71% Mismatches: 1
XX Query Watch: 92.56% Indels: 1
XX DB: 10 Gaps: 0
XX
XX US-10-763-972-2 (1-360) x ADC25999 (1-1017)
XX
XX Qy 1 MetLeuSerLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
XX Db 1 ATGCTGTCATTTTTCCTCCCTCAGGGGAGCAGCAGCGGAGCGCGCTGAGGCTCTG 60
XX
XX Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
XX Db 61 CTCCTGGAGGAGGAGCTCCCGGACATGGAGAGGTGACATGATATACATCAGGACAA 120
XX
XX Qy 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60

```

```

Db 121 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCGCTACAGTATC 180
Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
Db 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTCTCTGTGGCACTCTCTGGGGCAAAACC 240
Qy 81 LysArgTrpSerCysAlaThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGTGAACCTAGTGTGGCGGACCTGCTTTAT 300
Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpPropheGlyGlu 120
Db 301 GTGCTATTGCCCTTCTCCATCATCACCTACTACTAGATGACAGTGCCCTTTCGGGAG 360
Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 361 CTGCTCTGCAGCTGGTGCACTTCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGCT 420
Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 421 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGCCACCACTGTGTCTGCTGCC 480
Qy 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTroAlaLeuValValLeu 180
Db 481 TACCGACCCGACGATGCTGCTGGCGCACAGCACCACTCTGGGCCCTGGTGGTCTCTC 540
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
Db 541 CAGCTGTCTGCCACACTTGGCTTCTCCACACGGACTACATCAATGGCCAGATGATCTGG 600
Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 601 TATGACATGACGACGACCAAGAAATTTGATCGGCTTTTGTCTACGGCATAGTCTTGACA 660
Qy 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db 661 TTGCTGTGGCTTTT-TCCTCTCTGGTCAATTTGGTGTGTATTCACTGATGGTCAGGAG 719
Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db 720 CCTGATCAAGCCAGCAGGAGAACCTCATGAGGACAGGCAACACAGCCCGGACGATCAT 779
Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
Db 780 CCGGACCACTCTACTGGGTGTGGCTTTTCACTCTCTGTTTGTGGCTTCCATATCAC 839
Qy 281 SerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
Db 840 TCGCTCTCTTCTACCTACCACTCTGCTTCTGCTTCTTCTCAGGACTGCCAGCTCTTGATGGC 899
Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
Db 900 AGCCAGTGTGGCTTACAAGATATGGAGGCTCTGTTGAGTGTGACACCTGCTCAACCC 959
Qy 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeu 339
Db 960 AGTCTCTGTACTTTCTTCAAGGGGGCAAAATAAGTACAGTCAAGTCTCTCCAGAAACTG 1016
XX
XX RESULT 10
XX AAD61648
XX ID AAD61648 standard; DNA; 1017 BP.
XX AC AAD61648;
XX XX
XX DT 15-JAN-2004 (first entry)
XX DE Human GPCRx6 DNA.
XX
XX Human; G-protein coupled receptor; GPCR; infection; neoplastic process;
XX inflammation; myocardial infarction; atherosclerosis; angina pectoris;
XX hypertension; osteoporosis; antibacterial; cytostatic; fungicide; pain;
XX diabetes; cancer; virucide; analgesic; cardiant; gene; ds.

```

XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT 1..1017
 XX FT /*tag= a
 XX FT /product= "Human GPCR protein"
 XX
 XX PN US2003108986-A1.
 XX PD 12-JUN-2003.
 XX PF 20-FEB-2002; 2002US-00079384.
 XX XX
 XX PR 21-JUN-2001; 2001US-00885453.
 XX XX
 XX FA (EURO-) EUROSCREEN SA.
 XX XX
 XX PI Communi D, Lannoy V, Brezillon S, Datheux M, Parmentier M;
 XX PI Govaerts C;
 XX XX
 XX DR WPI; 2003-810852/76.
 XX DR P-ESDB; ABW00804.
 XX XX
 XX PT Novel G-protein coupled receptor useful for treating viral infections,
 XX PT bacterial infections, fungal infections, cancer, diabetes, hypertension,
 XX PT osteoporosis, angina pectoris, myocardial infarction, atherosclerosis.
 XX XX
 XX PS Claim 5; Fig 4; Opp; English.
 XX XX
 XX CC The present invention relates to novel G-protein coupled receptors
 XX CC (GPCRs) and the nucleic acids encoding them. The invention is useful for
 XX CC treating viral, bacterial and fungal infections, inflammatory and
 XX CC neoplastic processes, pain, diabetes, hypertension, osteoporosis, cancer,
 XX CC angina pectoris, myocardial infarction and atherosclerosis. The present
 XX CC sequence is human G-protein coupled receptor (GPCR) DNA
 XX
 XX SQ Sequence 1017 BP; 200 A; 306 C; 252 G; 259 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 7 68e-163 Length: 1017
 Score: 1792.00 Matches: 338
 Percent Similarity: 99.71% Conservative: 0
 Best Local Similarity: 99.71% Mismatches: 1
 Query Match: 92.56% Indels: 1
 DB: 10 Gaps: 0
 US-10-763-972-2 (1-360) x AAD61648 (1-1017)
 QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
 Db 1 ATGCTGTCCATTTCCTTCCTCAGGGGAAGCAGAGCGGAGCGCTCGTGAGCTCTG 60
 QY 21 LeuLeuGluGlyAlaSerArgCaspMetGluLysValAspMetAsnThrSerGlnGlu 40
 Db 61 CTCCTGAGGAGGCTCCCGGACATGGAGAGTGGACATGAATACATACAGGAACAA 120
 QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerile 60
 Db 121 GGTCTCTGCCAGTTCTCAGAGAGTACACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
 QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheSerTrpGlyGlnThr 80
 Db 181 ATCTTTATCTAGGGCTGCCACTAAATGGACCTGTCTTGGGCACCTCTGGGGCCAAACC 240
 QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
 Db 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGAACCTGATGTGTGGCGGACCTGCTTTAT 300
 QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpPheGlyGlu 120
 Db 301 GTGCTATTGCCCTTCCTTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 360

QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
 Db 361 CTGCTCTGCAAGCTGGTGCACTTCCTGTTCATATCAACCTTTTACGGCAGATCCTGCTG 420
 QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 Db 421 CTGACCTGCACTCTGTGCAACAGTTCCTAGTGTGTGCCACCCACTGTGTTCCTGCCC 480
 QY 161 TyrArgThrArgArgHisAlaThrLeuGlyThrSerThrThrThrAlaLeuValValLeu 180
 Db 481 TACCGGACCCGCGAGCATGCTGGCTGGGCACCAACACCACTGGGCCCTGGTGTCTCTC 540
 QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
 Db 541 CAGCTGTGCTGCCACACTGGCCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG 600
 QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
 Db 601 TATGACATGACCAAGCAAGAGAAATTTTATGATCGCTTTTGGCTACGGCATAGTTCTGACA 660
 QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
 Db 661 TTGCTGGCTTTT-TCCCTCCTTGGTCATTTTGGTGTCTATTCTACTGATGGTCAGGAG 719
 QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
 Db 720 CTTGATCAGCCAGAGAGAACCTTCATGAGACAGGCAACACACAGCCCGAGCGAGTCCAT 779
 QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
 Db 780 CCGGACCATCTACTGTGTGGCTCTTCACTCTCTGTGTGTGGCTTCCATATCATC 839
 QY 281 SerLeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
 Db 840 TCGCTCCTTCTACCTACCACTCTGCTTCTCTCTCTCAGGACTGCCAGCTCTTGATGGC 899
 QY 301 SerGlnCysGlyLeuGlnAspMetGluAspSerGlyGluCysGluGlnLeuProGlnPro 320
 Db 900 AGCCAGTGTGGCTTACAGATATGAGAGGCTCTCGTGAGTGTGAGCAGCTGCTCAACCC 959
 QY 321 SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeu 339
 Db 960 AGTCCTGTACTTCTTTCAGGGGGGCAAAATAAGATGAGTCTCAGGCTCTCCAGAAACTG 1016
 RESULT 11
 ID ABQ79300 standard; DNA; 1020 BP.
 XX AC ABQ79300;
 XX DT 17-OCT-2002 (first entry)
 XX DE Human GPCR designated PFI-020' encoding sequence.
 XX KW Human; GPCR; G-protein coupled receptor; antidepressant; neuroleptic;
 XX KW gene therapy; therapeutic; mood; depression; arousal; eating; sleeping;
 XX KW disorder; PFI-020'; gene; ds.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX FT 1..1020
 XX FT /*tag= a
 XX FT /product= "GPCR designated PFI-020"
 XX
 XX PN EP1215214-A1.
 XX PD 19-JUN-2002.
 XX PF 04-DEC-2001; 2001EP-00310137.
 XX PR 18-DEC-2000; 2000GB-00030855.
 XX PR 17-JAN-2001; 2001GB-00001222.
 PR

XX (PFIZ) PFIZER LTD.
PA (PFIZ) PFIZER INC.
XX
XX Fidoack MD;
XX
XX WPI; 2002-510798/55.
DR P-PSDB; ABB98146.
XX
XX New polynucleotide encoding G protein-coupled receptor PFI-020, useful
PT e.g. for treating eating and sleeping disorders and for identifying
PT specific modulators.
XX
XX Claim 1 (d); Page 12; 23pp; English.
XX
XX The invention relates to an isolated polynucleotide encoding a novel
CC polypeptide belonging to the class of proteins known as G-protein coupled
CC receptors (GPCRs). The activity of proteins of the invention may be
CC described as, antidepressant and neuroleptic. Polynucleotides of the
CC invention are used for recombinant expression of the G protein-coupled
CC receptor (PFI-020) polypeptides, to create transgenic animals, as source
CC of primers, probes, antisense sequences and ribozymes and in gene
CC therapy. Therapeutic agents of the invention can be used to treat a wide
CC range of disorders, particularly mood disorders, depression or arousal,
CC especially eating and sleeping disorders. The current sequence represents
CC a coding sequence for a human GPCR designated PFI-020.
XX
XX SQ Sequence 1020 BP; 200 A; 309 C; 252 G; 259 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,65e-162 Length: 1020
Score: 1785.00 Matches: 337
Percent Similarity: 99.12% Conservative: 1
Best Local Similarity: 98.83% Mismatches: 1
Query Match: 92.20% Indels: 2
6 Gaps: 0

US-10-763-972-2 (1-360) x ABQ79300 (1-1020)
QY 1 MetLeuSerIleLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
Db 1 ATGCTGCCATTTTGTCTCTCCAGGGAAGCAGAGCGGGAGCGTCTGGAGCTCTG 60
QY 21 LeuLeuGluGlyAlaSerArgMetGluLeuValAspMetAsnThrSerGlnGluGln 40
Db 61 CTCTCTGAGGAGCCCTCCGGGACATGGAGAGGTGACATGAATACATCAGAGAACAA 120
QY 41 GlyLeuCysGlnPheSerGluLeuValGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db 121 GGTCTGCGCAGTTCTCAGAGAGTACAGCAAGTCTACTCTCCCTGGCCCTACAGTATC 180
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
Db 181 ATCTTTATCTAGGGCTGCCACTAAATGGCAGCTGTCTTGTGGCAGCTCTCTGGGGCCAAACC 240
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 241 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCACCTGCTTTAT 300
QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120
Db 301 GTGCTATTGCCCTCTCATCATCACTACTACTAGTACAGGTTGGCCCTTCGGGGAG 360
QY 121 LeuLeuCysIleLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 361 CTGCTCTGCAAGTGGTGGCACTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 421 CTGACCTGCATCTCTGTCACCACTCTAGTGTGGCAGCCACTGTTGCTGCTGCC 480
QY 161 TyrArgThrArgGHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180

Db 481 TACCGGACCGGAGGATGCTGGTGGGACACGACACACCTGGGGCCCTGGTGGTCTC 540
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
Db 541 CAGTGTGCTGCCACACATGGGCTTCTCCACAGGACTACATCAATGCGCAGATGATCG 600
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 601 TATGACATGACCGACCAAGAGATTTTGTATCGGCTTTTGGCTACGGCATAGTCTTGACA 660
QY 221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG 240
Db 661 TTGTCTGGCTTTCTTCTCCCTCTCTTGGTGTATTTGGTGTGCTATTCACGATGGTCA 720
QY 240 LuProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH 260
Db 721 AGCTGTATCAGCCAGAGGAGACCTCATGAGGACAGGACACAGCCGAGCCAGGTCC 780
QY 260 isProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrH 280
Db 781 ATCCGACCATCTCTACTGTTGTGGCTCTCTCACCCCTCTGTTGTGGCTCTCCATATC 840
QY 280 isSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
Db 841 ACTGCTCTCTCTACCTACCACTGCTTCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGA 900
QY 300 LysSerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320
Db 901 GCACCCAGTGTGGCTCAAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAAC 960
QY 320 roSerProValLeuSerPheIleGlyValAsnArgValAtcLeuLeuGlnLysLeu 339
Db 961 CCAGTCTCTGTTACTTCTTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCTCCAGAACTG 1019

RESULT 12
AAD29667
ID AAD29667 standard; cDNA; 1076 BP.
XX
AC AAD29667;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human G-protein coupled receptor (GCRC-1) cDNA.
KW Human; G-protein coupled receptor; GCRC-1; cell proliferative disorder;
KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;
KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;
KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;
KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;
KW neurotropic; neuroprotective; cardiant; immunosuppressive; anorectic;
KW virucide; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 57..1076
FT /*tag= a
FT /product= "Human GCRC-1"
XX
XX WC00210387-A2.
XX
XX 07-FEB-2002.
XX
XX 25-JUL-2001; 2001WO-US023433.
XX
XX 27-JUL-2000; 2000US-0221478P.
PR 03-AUG-2000; 2000US-0223268P.
PR 21-AUG-2000; 2000US-0227054P.
PR 08-SEP-2000; 2000US-0231121P.
PR 13-SEP-2000; 2000US-0232243P.

PR 15-SEP-2000; 2000US-0232691P.
 PR 22-SEP-2000; 2000US-0235146P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;
 PI Elliot VS, Ramkumar J, Kailick DA, Walia NK, Hafalia AJA;
 PI Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L, Graul RC;
 PI Warren BA, Lee EA, Ding L;
 XX WPI; 2002-188744/24.
 DR P-PSDB; AAE18640.
 XX
 XX New human G-protein coupled receptor polypeptide for diagnosis,
 PT prevention and treatment of cell proliferative, neurological,
 PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic
 PT disorders.
 XX
 PS Claim 5; Page 137-138; 150pp; English.
 XX
 CC The invention relates to novel human G-protein coupled receptors (GPCR)
 CC and their encoding polynucleotides. GPCR is useful as an immunogen for
 CC preparing monoclonal and polyclonal antibodies. GPCR is useful for
 CC diagnosing, treating and preventing a cell proliferative disorder (e.g.,
 CC hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,
 CC Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's
 CC disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,
 CC myocardial infarction), gastrointestinal disorder (e.g., gastritis,
 CC cirrhosis, Crohn's disease), an autoimmune/inflammatory disorder (e.g.,
 CC acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,
 CC rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,
 CC osteoporosis), and viral infections. GPCR is useful in a number of drug
 CC screening techniques, and to analyse the proteome of a tissue or cell
 CC type. GPCR is useful for creating knockin humanised animals or
 CC transgenic animals to model human diseases, in somatic or germline gene
 CC therapy, to generate a transcript image of a tissue or cell type, for
 CC detecting differences in the chromosomal location due to translocation,
 CC inversion, etc., among normal, carrier or affected individuals, and as
 CC hybridization probes for mapping naturally occurring genomic sequences.
 CC GPCR is useful in Southern or northern analysis, dot blot or other
 CC membrane-based technologies, in PCR technologies, in dipstick, pin,
 CC multi-format enzyme linked immunosorbent (ELISA)-like assays, and in
 CC microarrays utilising fluids or tissues from patients to detect altered
 CC GPCR expression. The present sequence is human GPCR-1 cDNA
 XX
 SQ Sequence 1076 BP; 211 A; 318 C; 274 G; 273 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.94e-162 Length: 1076
 Score: 1785.00 Matches: 337
 Percent Similarity: 99.12% Conservative: 1
 Best Local Similarity: 98.83% Mismatches: 1
 Query Match: 92.20% Indels: 2
 DB: 6 Gaps: 0

US-10-763-972-2 (1-360) x RAD29667 (1-1076)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
 Db 57 ATGCTGCTCCATTTCCTTCCTCCAGGGGAAGCAGAGGGGAGCGCTCGTGAGCTCTG 116

Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLeuValAspMetAsnThrSerGlnGluIn 40
 Db 117 CTCTTGAGGAGGACCTCCCGGACATGGAGAGGTGGACATGAATCATCATCAGGAACAA 176

Qy 41 GlyLeuGlnPheSerGluLeuValLeuValLeuValLeuValLeuValLeuValLeu 60
 Db 177 GGTCTCTGCCAGTCTCAGAGAGTACACAGAGTCTACCTCTCTCTGCTTACAGTATC 236

Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheSerTrpGlyGlnThr 80
 Db 237 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCCAACC 296

Qy 81 LysArgTrpSerCysAlaThrThrTyrIleuValAsnLeuMetValAlaAspLeuLeuTyr 100
 Db 297 AAGCGCTGGAGCTGTGCCACACACCTATCTGGTGAACTGATGGTGGCGGACCTCTTTAT 356

Qy 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120
 Db 357 GTGCTATTGCCCTTCTCTCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 416

Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
 Db 417 CTGCTCTGCAAGCTGGTGCACTTCTCTTATATCAACCTTTACGGCAGCATCTCTGCTG 476

Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
 Db 477 CTGACCTGCATCTCTGTGCACCACTTCTTAGGTGTGGCACCCACATGTGTCTGCTGCC 536

Qy 161 TyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeu 180
 Db 537 TACCGACCCGCGCATGCTGCTGGCCACCAACACCACTGGGCTGGTGGTCTCTC 596

Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
 Db 597 CAGCTGTGCCACACCTGCTTCTCCACACCGGACTACATCAATGGCCAGATGATCTGG 656

Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
 Db 657 TATGACATGACACCGCAGAGAGAAATTTGATCGGCTTTTGCCTACGGCATAGTCTGACA 716

Qy 221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGln 240
 Db 717 TTGCTGCTGCTTCTTCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 776

Qy 240 LuProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValH 260
 Db 777 AGCTGTATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCGCGGAGCTCC 836

Qy 260 IsProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTrpH 280
 Db 837 ATCCGGACCATCTCTACTGCTGTGTGGCTCTTCACTCCCTCTGTTTGTGCTCTTCCATC 896

Qy 280 IsSerLeuLeuLeuProHisIleLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
 Db 897 ACTGCTCTCTTCTACCTACCATCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGATG 956

Qy 300 LysSerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnLeuProGlnP 320
 Db 957 GCACCCAGTGTGGCTTACAGATATGAGGAGCTCTGCTGAGTGTGAGCAGCTGCTCAAC 1016

Qy 320 roSerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeu 339
 Db 1017 CCAGTCTGTACTTTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCTCAGAACTG 1075

RESULT 13
 ADO30395
 ID ADO30395 standard; cDNA; 1076 BP.
 XX
 AC ADO30395;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human GPCR P2Y3L polynucleotide, SEQ ID NO:1498.
 XX
 KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
 KW cytosstatic; antinflammatory; vasotropic; antiarrhythmic;
 KW CNS; central nervous system; respiratory; antiarrhythmic; antidiabetic;

XX AAS17747;
 AC 26-FEB-2002 (first entry)
 DT Human P2Y-like G protein-coupled receptor exon 1.
 DE
 XX Human; ds; P2Y-like G protein-coupled receptor; GPCR; COPD;
 KW chronic obstructive pulmonary disease; nervous system disease;
 KW Parkinson's disease; multiple sclerosis; dementia; stroke;
 KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;
 KW bacterial infection; fungal infection; protozoan infection;
 KW viral infection; pain; cancer; anorexia; bulimia; asthma; exon 1;
 KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;
 KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;
 KW psychotic disorder; neurological disorder; anxiety; schizophrenia;
 KW manic depression; delirium; severe mental retardation; dyskinesia.
 XX Homo sapiens.
 OS
 XX WO2001:85764-A2.
 EN
 XX 15-NOV-2001.
 PD
 XX 09-MAY-2001; 2001WO-EP005244.
 PF
 XX 11-MAY-2000; 2000US-0203582P.
 PR
 XX 21-FEB-2001; 2001US-0265857P.
 PA (FARB) BAYER AG.
 PI Ramakrishnan S;
 XX WPI; 2002-075242/10.
 DR
 XX New polynucleotides for producing P2Y-like G protein-coupled receptors
 PT (GPCR) that are used for screening inhibitors or regulators of human P2Y-
 PT like GPCR, especially useful for treating pain, cancer or neurological
 PT disorders.
 XX Claim 1; Fig 4; 11app; English.
 PS
 XX The invention relates to an isolated polynucleotide encoding a P2Y-like G
 CC protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or
 CC allele, a host cell containing an expression vector comprising the
 CC polynucleotide and screening for agents that regulate the GPCR activity.
 CC The polynucleotide is useful for producing P2Y-like GPCR polypeptide,
 CC which may be employed for screening agents that inhibit or regulate human
 CC P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is
 CC useful for treating or ameliorating P2Y-like GPCR disorders, particularly
 CC COPD (chronic obstructive pulmonary disease), peripheral or central
 CC nervous system disease (e.g. Parkinson's disease, multiple sclerosis,
 CC dementia, stroke, Alzheimer's disease and many other diseases and
 CC disorders listed in the specification), benign prostatic hyperplasia or
 CC urinary incontinence. A pharmaceutical composition containing the
 CC modulators and/or regulators of P2Y-like GPCR is useful for modulating
 CC the activity of a P2Y-like GPCR. In particular, these are useful for
 CC treating, preventing or ameliorating infections (e.g. bacterial, fungal,
 CC protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma,
 CC acute heart failure, hypotension, hypertension, osteoporosis, diabetes,
 CC angina pectoris, myocardial infarction, ulcers, inflammation, allergies,
 CC psychotic or neurological disorders (e.g. anxiety, schizophrenia, manic
 CC depression, delirium, severe mental retardation or dyskinesias). The
 CC present sequence is exon 1 of the P2Y-like GPCR of the invention
 XX
 SQ Sequence 850 BP; 165 A; 262 C; 201 G; 222 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9,46e-139 Length: 850
 Score: 1541.00 Matches: 283
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 79.60% Indels: 0

DB: 6 Gaps: 0
 US-10-763-972-2 (1-360) x AAS17747 (1-850)
 QY 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeuCysGlnPheSerGluLys 48
 DB 1 ATGGAGAGAGTGGACATGATACATACAGGAACAAGGTCTCTGCCAGATTCTCAGAGAAG 60
 QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerLeilellePheLeuGlyLeuProLeu 68
 DB 61 TACAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTTATCCAGGGCTGCCACTA 120
 QY 69 AsnGlyThrValLeuTyrHisSerTyrGlyGlnThrLysArgTyrPsrCysAlaThrThr 88
 DB 121 AATGGCACTGTCTGTGGCACTCTCTGGGGCAACCAAGCGCTGGAGCTGTGCCACCACC 180
 QY 89 TyrLeuValAsnLeuMetValAlaAspLeuTyrValLeuLeuProPheLeuLeilelle 108
 DB 181 TATCTGTGAACCTGATGGTGGCGAGCTGCTTTATGTGCTATTGCGCTTCTCATCATC 240
 QY 109 ThrTyrSerLeuAspAspArgTyrProPheGlyGluLeuLeuCysLysLeuValHisPhe 128
 DB 241 ACCTACTACTAGATGACAGGTGGCCCTTCGGGAGGTGCTCTGCAAGCTGTGTCATTC 300
 QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148
 DB 301 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCATCTCTGTGCACAG 360
 QY 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgGlnHisAlaTyr 168
 DB 361 TTCTAGTGTGTGCCACCCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 169 LeuGlyThrSerThrThrTyrAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188
 DB 421 CTGGCACACGACACCACTGGGGCTGGTGGTCTCCAGCTGCTGCCACACTGGCTTC 480
 QY 189 SerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGlnGluAsn 208
 DB 481 TCCACACGAGTACTACATCAATGGCCAGATGATGATGATGATGATGATGATGATGAT 540
 QY 209 PheAspArgLeuPheAlaTyrGlyLeuValLeuThrLeuSerGlyPheLeuSerLeuLeu 228
 DB 541 TTTGATCGCTTTTGGCTACGGCATAGTCTGACATGCTGCTGCTTCTTCCCTCTCT 600
 QY 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248
 DB 601 GGTCAATTTGGTGTCTATTCACTGATGCTCAGGAGCTGATCAAGCCAGAGGAGAACCT 660
 QY 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTyr 268
 DB 661 CATGAGGACAGGCAACACAGCCCGAGCGAGTCCATCCGGACCATCTTACTGTGTGTGG 720
 QY 269 ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu 288
 DB 721 CCTCTTCACTGCTTTTGGTGGCTTCCATCATCTGCTGCTTCTTACTCATCATCTG 780
 QY 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
 DB 781 CTTTCTGCTTTCTCAGGACTCCAGCTCTTGTATGATGCGCCAGTGTGCGCTTACAGATATG 840
 QY 309 GluAlaSer 311
 DB 841 GAGGCTCT 849
 RESULT 15
 ID ADO30397
 AC ADO30397;
 XX
 DT 29-JUL-2004 (first entry)
 XX Mouse GPCR P2Y3L polynucleotide, SEQ ID NO:1500.
 DE

XX	G protein-coupled receptor; GPCR; drug screening; diagnosis;	CC	ftp.wipo.int/pub/published_pct_sequences.
KW	transgenic mouse; neurological disorder; adrenal gland disorder;	XX	
KW	colon disorder; intestinal disorder; cardiovascular disorder;	SQ	Sequence 906 BP; 173 A; 276 C; 216 G; 241 T; 0 U; 0 Other;
KW	muscular disorder; blood disorder; immune disorder; bone disorder;		
KW	joint disorder; metabolic disorder; nutritive disorder; cancer;	Alignment Scores:	
KW	kidney disorder; liver disorder; lung disorder; breast disorder;	Pred. No.:	1.14e-92 — Length: 906
KW	ovary disorder; uterus disorder; prostate disorder; testis disorder;	Score:	1063.50 Matches: 207
KW	skin disorder; stomach disorder; pancreas disorder; spleen disorder;	Percent Similarity:	77.23% Conservative: 27
KW	thymus disorder; thyroid disorder; antiparkinsonian; antimanic;	Best Local Similarity:	68.32% Mismatches: 66
KW	cytostatic; antiinflammatory; vasotropic; antiangiinal; antiarrhythmic;	Query Match:	54.93% Indels: 3
KW	CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;	DB:	12 Gaps: 1
KW	virulence; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;		
KW	dermatological; antiulcer; antithyroid; antiallergic; anorectic;	US-10-763-972-2 (1-360) x ADO30397 (1-906)	
KW	immunosuppressive; nephrotropic; gene therapy; GPCR modulator; mouse;		
KW	murine; gene; ss.		
XX	Mus musculus.		
XX			
PN	W02004040000-A2.		
XX			
PD	13-MAY-2004.		
XX			
XX	09-SEP-2003; 2003WO-US028226.		
PF			
XX			
XX	09-SEP-2002; 2002US-0409303P.		
PR			
XX	09-APR-2003; 2003US-0461329P.		
XX			
XX	(PRIM-) PRIMAL INC.		
PA			
XX			
PI	Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;		
PI	Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;		
XX			
DR	WPI; 2004-390329/36.		
DR	P-PSDB; ADO30396.		
XX			
XX	Novel mammalian G protein coupled receptors, useful for identifying		
PT	compounds that modulates diagnosing and treating disease condition		
PT	associated with GPCR dysfunction e.g. autoimmune diseases, angina		
PT	pectoris, Parkinson's disease.		
XX			
XX	Claim 151; SEQ ID NO 1500; 542pp; English.		
PS			
XX			
CC	The invention relates to human and mouse G protein-coupled receptors		
CC	(GPCRs) and nucleic acids encoding them. The invention also relates to		
CC	sequences at least 90% identical to the GPCR proteins and nucleic acids		
CC	of the invention; methods of treating, preventing or diagnosing diseases		
CC	associated with GPCRs of the invention; methods of screening for		
CC	compounds useful in the treatment of GPCR-related diseases; a transgenic		
CC	mouse comprising a GPCR gene of the invention; a mouse comprising a		
CC	mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived		
CC	from the transgenic mice; kits comprising several mice, each of which has		
CC	a mutation in a different GPCR gene of the invention; and kits comprising		
CC	probes which hybridise to GPCR polynucleotides of the invention. The		
CC	invention further discloses variants of the GPCR polypeptides and vectors		
CC	comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may		
CC	be used in the diagnosis, treatment or prevention of a wide variety of		
CC	diseases including neurological disorders (e.g., Alzheimer's disease,		
CC	depression, diabetic neuropathy, Parkinson's disease or schizophrenia);		
CC	disorders of the adrenal gland; disorders of the colon or intestine		
CC	(e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel		
CC	syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or		
CC	myocardial infarction); muscular disorders; blood disorders (e.g.,		
CC	anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or		
CC	AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid		
CC	arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,		
CC	obesity, enzyme deficiency-related diseases or vitamin deficiency-related		
CC	diseases); and disorders of the kidney, liver, lung, breast, ovary,		
CC	uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and		
CC	thyroid (e.g., cancers). The present sequence represents a GPCR-encoding		
CC	nucleic acid of the invention. Note: The full sequence data for this		
CC	patent did not form part of the printed specification; those sequences		
CC	not shown were obtained in electronic format directly from WIPO at		

Qy 327 ysgly 328
Db 901 CAGGG 905

Search completed: November 10, 2004, 12:11:41
Job time : 602 secs

10/7639772
Seq. IDs 1 & 2

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 08:36:32 ; Search time 4772 Seconds
(without alignments)
10722.434 Million cell updates/sec

Title: US-10-763-972-1

Perfect score: 1082
Sequence: 1 atgctgtccattttgttccc.....ccagggttgacagatctgg 1082

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 segs, 2364489745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_cm.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_br.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1082	100.0	1082	6	BD187416
2	1082	100.0	1082	6	AX458306
3	1082	100.0	1145	9	HS345013
4	1082	100.0	1334	6	AX646427
5	1082	100.0	1334	9	AB056670
6	1082	100.0	3143	6	AX385511
7	1082	100.0	164502	2	AC026960
8	1082	100.0	167084	9	AC092999
9	1071	99.0	163958	2	AC021773
10	1059	97.9	1108	6	AX921825
11	1007	93.1	1019	9	AF411108
12	1002.8	92.7	1020	6	BD187417
13	1002.8	92.7	1020	6	AX458308
14	1002.8	92.7	1076	6	AX375230
15	996.4	92.1	1002	6	BD095704
16	850	78.6	850	6	AX385514
17	779.4	72.0	1140	6	CQ734310
18	649.4	60.0	184377	10	AC121579
19	649.4	60.0	195290	2	AC135635

20	644.4	59.6	250610	2	AC120742
21	644.4	59.6	276127	2	AC116265
22	510	47.1	510	6	AX147768
23	510	47.1	510	6	AX521817
24	396.4	36.6	681	6	AX244718
25	370	34.2	585	6	AX244727
26	235.8	21.8	1163	5	GD2733
27	227.8	21.1	1074	5	AF069555
28	227.8	21.1	2025	6	AR270919
29	227.8	21.1	2025	6	AX548931
30	227.8	21.1	2025	9	HSU07225
31	226.2	20.9	2257	9	BC028135
32	224.6	20.8	1134	9	AX136753
33	224.6	20.8	1974	6	CQ720945
34	224.6	20.8	2681	9	BC012104
35	224.6	20.8	123185	2	AF000587
36	224.6	20.8	196988	9	AF002761
37	222	20.5	551	6	AX230145
38	203.8	18.8	1842	6	I34573
39	203.8	18.8	1842	6	I36524
40	203.8	18.8	1842	6	I76280
41	202.2	18.7	2138	6	AX454153
42	202.2	18.7	2138	10	MUSP2UREC
43	202.2	18.7	2562	10	BC006613
44	201.4	18.6	1977	5	AF031897
45	200.6	18.5	175947	2	AC079641

ALIGNMENTS

RESULT 1	BD187416	BD187416	1082 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD187416	Novel polypeptide.				
DEFINITION	BD187416					
ACCESSION	BD187416.1	GI:32997155				
VERSION	JP 2003009885-A/1.					
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 1082)					
AUTHORS	Fidock,M.D.					
TITLE	Novel polypeptide					
JOURNAL	Patent: JP 2003009885-A 1 14-JAN-2003;					
COMMENT	Pfizer Ltd (BP/GB) only; Pfizer Inc (US JP EP except GB)					
	OS homo sapiens					
	PN JP 2003009885-A/1					
	PD 14-JAN-2003					
	PF 17-DEC-2001 JP 2001382707					
	PR 18-DEC-2000 GB 0030855.1,17-JAN-2001 GB 0101222.8 PI					
	mark david fidock					
	CC					

FEATURES	source	Location/Qualifiers
		1..1082
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
ORIGIN		
Query Match	100.0%;	Score 1082; DB 6; Length 1082;
Best Local Similarity	100.0%;	Pred. No. 6.7e-267;
Matches 1082;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	ATGCTGTCCATTTTCCTTCCTCCGGGAGACAGACGGGAGCCGCGTGGAGCTCTG 60
DB	1	ATGCTGTCCATTTTCCTTCCTCCGGGAGACAGACGGGAGCCGCGTGGAGCTCTG 60
QY	61	CTCCTGGAGGAGCCTCCCGGACATGAGAGGTGGACATGATACATCAGGAACAA 120
DB	61	CTCCTGGAGGAGCCTCCCGGACATGAGAGGTGGACATGATACATCAGGAACAA 120

Qy	121	GGTCTCTGCCAGTTCTCAGAGAAAGTCAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC	180
Db	121	GGTCTCTGCCAGTTCTCAGAGAAAGTCAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC	180
Qy	181	ATCTTTATCCTTAGGGTCCACTAAATGGGCAGTCTGTGTGGCACTCTCTGGGCGCCAAACC	240
Db	181	ATCTTTATCCTTAGGGTCCACTAAATGGGCAGTCTGTGTGGCACTCTCTGGGCGCCAAACC	240
Qy	241	AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACTGTGTGGCGCGACCTGCTTTAT	300
Db	241	AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACTGTGTGGCGCGACCTGCTTTAT	300
Qy	301	GTGCTATTGCCCTTCCTCATCATCACTACTACTAGATGACACAGTGGCCCTTGGGGGAG	360
Db	301	GTGCTATTGCCCTTCCTCATCATCACTACTACTAGATGACACAGTGGCCCTTGGGGGAG	360
Qy	361	CTGCTCTGCAAGCTGGTGCACTTCTCTGTATTATCAACCTTTACGGCAGCATCTGCTG	420
Db	361	CTGCTCTGCAAGCTGGTGCACTTCTCTGTATTATCAACCTTTACGGCAGCATCTGCTG	420
Qy	421	CTGACCTGCAATCTCTGTGCAACGATTCCTTAGTGTGTGCCACCACTGTGTTTGGTGGCC	480
Db	421	CTGACCTGCAATCTCTGTGCAACGATTCCTTAGTGTGTGCCACCACTGTGTTTGGTGGCC	480
Qy	481	TACCGGACCCGACGGCATGCTGGCTGGGCAACAGCACCACTGGGGCCCTGGTGGTCTC	540
Db	481	TACCGGACCCGACGGCATGCTGGCTGGGCAACAGCACCACTGGGGCCCTGGTGGTCTC	540
Qy	541	CAGCTGTGCCCACTGTGGCTTCTCCCAACCGGACTACATCAATGGCCAGATGATCTGG	600
Db	541	CAGCTGTGCCCACTGTGGCTTCTCCCAACCGGACTACATCAATGGCCAGATGATCTGG	600
Qy	601	TATGACATGACCCAGCAAGAGATTTTGTGCGCTTTTCCCTACGGCAGTAGTTCTGACA	660
Db	601	TATGACATGACCCAGCAAGAGATTTTGTGCGCTTTTCCCTACGGCAGTAGTTCTGACA	660
Qy	661	TTGTCCTGGCTTTCTTTCCCTCCTTGGTCAATTTTGGTGTCTATTCACTGATGTGCAGGAG	720
Db	661	TTGTCCTGGCTTTCTTTCCCTCCTTGGTCAATTTTGGTGTCTATTCACTGATGTGCAGGAG	720
Qy	721	CCTGATCAAGCAGAGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCCAGGTCCAT	780
Db	721	CCTGATCAAGCAGAGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCCAGGTCCAT	780
Qy	781	CCGGACCATCCTACTGTGTGTGGCCCTTTCACCCCTCTGTTTGTGGCCCTTCCATATCAC	840
Db	781	CCGGACCATCCTACTGTGTGTGGCCCTTTCACCCCTCTGTTTGTGGCCCTTCCATATCAC	840
Qy	841	TGCGTCTCTTCACTCAACATCTGCTTTCTGCTTCTCAGACCTGCCAGCTCTTGATGGC	900
Db	841	TGCGTCTCTTCACTCAACATCTGCTTTCTGCTTCTCAGACCTGCCAGCTCTTGATGGC	900
Qy	901	AGCCAGTGTGGCCCTTACAAGATATGGAGGCCCTCTGGTGAGTGTGAGCAGCTGCTCAACCC	960
Db	901	AGCCAGTGTGGCCCTTACAAGATATGGAGGCCCTCTGGTGAGTGTGAGCAGCTGCTCAACCC	960
Qy	961	AGTCTCTGATCTTTTCAAGGGGGGCAAAAATAGAGTCAAGTCTCTTCCAGAACTGAGG	1020
Db	961	AGTCTCTGATCTTTTCAAGGGGGGCAAAAATAGAGTCAAGTCTCTTCCAGAACTGAGG	1020
Qy	1021	CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCCAAGGTTGAACAGATCT	1080
Db	1021	CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCCAAGGTTGAACAGATCT	1080
Qy	1081	GG 1082	
Db	1081	GG 1082	

RESULT 2
AX458306
LOCUS

AX458306 1082 bp DNA linear PAT 08-JUL-2002

721 CCTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCAGGTCCTAT 780
781 CCGGACATCCTACTGGTGTGGGCTCTTACCCCTCTGTTTGTGCGCTTCCATATCAC 840
781 CCGGACATCCTACTGGTGTGGGCTCTTACCCCTCTGTTTGTGCGCTTCCATATCAC 840
841 TCGCTCCTTCTACTCTACCACTCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGC 900
841 TCGCTCCTTCTACTCTACCACTCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGC 900
901 AGCAGTGTGGCTACAGATATGAGGAGCTCTGGTGTGAGTGTGAGAGCTGCCATCAACC 960
901 AGCAGTGTGGCTACAGATATGAGGAGCTCTGGTGTGAGTGTGAGAGCTGCCATCAACC 960
961 AGTCTCTACTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCCAGAACTGAGG 1020
961 AGTCTCTACTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCCAGAACTGAGG 1020
1021 CAGAACAGTGTGGTGTGAGATCCAGCTGTGGAGAGAGATGCCAGGTTGAACAGATCT 1080
1021 CAGAACAGTGTGGTGTGAGATCCAGCTGTGGAGAGAGATGCCAGGTTGAACAGATCT 1080
1081 GG 1082
1081 GG 1082

RESULT 3
HSA345013 1145 bp DNA linear PRI 07-SEP-2003
LOCUS HSA345013
DEFINITION Homo sapiens P2Y2-like gene for nucleoside/nucleotide receptor.
ACCESSION AJ345013
VERSION AJ345013.1 GI:34495183
KEYWORDS nucleoside/nucleotide receptor; P2Y2-like gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Bruss, M., Bonisch, H. and Kugelgen, I.
TITLE Molecular identification and functional characterization of a new G protein-coupled nucleoside-/nucleotide receptor
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1145)
AUTHORS Bruess, M.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2001) Bruess M., Pharmacology and Toxicology, University of Bonn, Reuterstrasse 2 b, D-53113, GERMANY

FEATURES
source
1..1145
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/tissue_type="blood"
10..1095
/gene="p2y2-like"
10..1095
/gene="P2Y2-like"
/function="G protein coupled receptor"
/codon_start=1
/product="nucleoside/nucleotide receptor"
/protein_id="CAC87811.1"
/db_xref="GI:34495184"
/translation="MLSILLPSRSGSRGALLLEGASRDMKVDNMTSQBGLCQ
FSKIKVYLUSLWYFIIFLPLNGVLVHWSGQTKRWSCATIYLVNLMVADLLYL
LPELIYVLDLDRMPFGLLCKLVHFLYINLYGSILLTFCISVHQLGVCPLSLP
YRTRHAWLGTSTWLVVQLLPTAFSTHDYINGQMIWYDNTSGNDFRLPAYGI
LTLSCFLSLHGFVFTDQEPDQARGEPEHQRQSPSOVHPDPTGVNPLHPLFCA
LPVHLLPPLLHLSAFSGLPALDQSCGLQDMEASGCEQLPQPSVLPSPKGGKVR
LLQLRQNLKEHPAGKRCFGLNRSG"

gene
CDS

ORIGIN

Query Match 100.0%; Score 1082; DB 9; Length 1145;
Best Local Similarity 100.0%; Pred. No. 6.7e-267;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCCATTTTGCTTCTCCTTCAGGGGAAGCAGAGCGGAGCCCTCGTGGAGCTCTG 60
DB 10 ATGCTGTCCATTTTGCTTCTCCTTCAGGGGAAGCAGAGCGGAGCCCTCGTGGAGCTCTG 69
QY 61 CTCCTGGAGGAGCCTCCCGGACATGGAGAAGTGGACATGAATATACATACAGAGAA 120
DB 70 CTCCTGGAGGAGCCTCCCGGACATGGAGAAGTGGACATGAATATACATACAGAGAA 129
QY 121 GGTCTCTGCCAGTCTCAGAGAAATCAAGCAAGTCTACCTCTCCCTGGGCCCTACAGTATC 180
DB 130 GGTCTCTGCCAGTCTCAGAGAAATCAAGCAAGTCTACCTCTCCCTGGGCCCTACAGTATC 189
QY 181 ATCTTTATCCCTAGGCTCCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGGCAAAACC 240
DB 190 ATCTTTATCCCTAGGCTCCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGGCAAAACC 249
QY 241 AAGCGCTGGAGCTGTGCCAACCACTATCTGGTGAACCTGTATGGTGGCGCACTGCTTTAT 300
DB 250 AAGCGCTGGAGCTGTGCCAACCACTATCTGGTGAACCTGTATGGTGGCGCACTGCTTTAT 309
QY 301 GTGCTATTGGCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGAG 360
DB 310 GTGCTATTGGCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGAG 369
QY 361 CTGCTCTCAAGCTGGTGCACTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
DB 370 CTGCTCTCAAGCTGGTGCACTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 429
QY 421 CTGACCTGATCTCTGTGCAACAGTCTTAGTGTGTGCCACCACTGCTGTGTGCTGCC 480
DB 430 CTGACCTGATCTCTGTGCAACAGTCTTAGTGTGTGCCACCACTGCTGTGTGCTGCC 489
QY 481 TACCGGACCCGAGGATGCTGCTGGTGGGCAACAGCACCCTGGGCCCTGGTGGCTCTC 540
DB 490 TACCGGACCCGAGGATGCTGCTGGTGGGCAACAGCACCCTGGGCCCTGGTGGCTCTC 549
QY 541 CAGCTGTGCCCACTGGCTTCTCCACAGGACTACATCAATATGAGGCAAGATATCTGG 600
DB 550 CAGCTGTGCCCACTGGCTTCTCCACAGGACTACATCAATATGAGGCAAGATATCTGG 609
QY 601 TATGACATGACAGCAGAGAGATTTTGTATCGGCTTTTGTGCTTACGCGCATGCTTGACA 660
DB 610 TATGACATGACAGCAGAGAGATTTTGTATCGGCTTTTGTGCTTACGCGCATGCTTGACA 669
QY 661 TTGCTGTGCTTTCTTTTCT 720
DB 670 TTGCTGTGCTTTCTTTTCT 729
QY 721 CTTGATCAAGCAGAGAGAACCTCATGAGGACAGGCAACACAGGCCGAGCAGGTCCAT 780
DB 730 CTTGATCAAGCAGAGAGAACCTCATGAGGACAGGCAACACAGGCCGAGCAGGTCCAT 789
QY 781 CCGGACCATCTTACTTGTGTGTGGGCTCTTACCCCTCTGTTTGTGGCCCTTCCATATCAC 840
DB 790 CCGGACCATCTTACTTGTGTGTGGGCTCTTACCCCTCTGTTTGTGGCCCTTCCATATCAC 849
QY 841 TCGCTCCTTCTACTCTACCACTCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGC 900
DB 850 TCGCTCCTTCTACTCTACCACTCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGC 909
QY 901 AGCCAGTGTGGCTACAGATATGAGGAGCTCTGGTGTGAGTGTGAGAGCTGCCATCAACC 960
DB 910 AGCCAGTGTGGCTACAGATATGAGGAGCTCTGGTGTGAGTGTGAGAGCTGCCATCAACC 969
QY 961 AGTCTCTACTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCCAGAACTGAGG 1020
DB 970 AGTCTCTACTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCCAGAACTGAGG 1029

```
QY 1021 CAGAACAGTTGGTGAGCATCCAGCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCT 1080
|||||
Db 1030 CAGAACAGTTGGTGAGCATCCAGCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCT 1089
|||||
QY 1081 GG 1082
|||
Db 1090 GG 1091
|||||

RESULT 4
AX646427 1334 bp DNA linear PAT 04-MAR-2003
LOCUS
DEFINITION Sequence 619 from Patent EP1270724.
ACCESSION AX646427
VERSION AX646427.1 GI:28798808
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Suwa,M., Asai,K., Akiyama,Y. and Aburatani,H.
TITLE Guanosine triphosphate-binding protein coupled receptors
JOURNAL Patent: EP 1270724-A 619 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JP) ; Center for Advanced Science and Technology Incubation, Ltd.
(JP)

FEATURES
source
1..1334
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
join(201..715,834..1134)
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD69342.1"
/db_xref="GI:28798809"
/translaton="MEKVDNMTSQEGLQCFSEKYQVLYSLAYSIIIFILGLPLNGTV
LMSWGOTKWSQATTVLNLVADLLVLLPLITYSLDDRPFGELLCKLVHPLF
YINLYGILLTCSVHQPICVCHPLCSLPYRFRANLGTSTNALVILQLPLAF
SHYDIYNGQMWISLKEEENLMTGNTPARSIRTLILVGLFTLCFVFPFHTRSYL
TICFLSQDCLLMAASVAYKIWRPLVSVSCLNPLVLYFLSRGAKTESGSSRN"

ORIGIN
Query Match 100.0%; Score 1082; DB 6; Length 1334;
Best Local Similarity 100.0%; Pred. No 6.8e-267;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCCATTTTGTCTTCCAGGGGAGCAGAGAGCGGCGCGTGTGAGCTCTG 60
Db 117 ATGCTGTCCATTTTGTCTTCCAGGGGAGCAGAGAGCGGCGCGTGTGAGCTCTG 176

QY 61 CTCCTGGAGGAGCTCCCGGACATCGAGAGGTGACATGATATACATCACAGGACAA 120
Db 177 CTCCTGGAGGAGCTCCCGGACATCGAGAGGTGACATGATATACATCACAGGACAA 236

QY 121 GGTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
Db 237 GGTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 296

QY 181 ATCTTTATCTTAGGCTGCCAATAATGGCACTGTCTTTGGCACTCTCTGGGCGCCAAACC 240
Db 297 ATCTTTATCTTAGGCTGCCAATAATGGCACTGTCTTTGGCACTCTCTGGGCGCCAAACC 356

QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTAACCTGATGTGGCGGACCTGCTTTAT 300
Db 357 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTAACCTGATGTGGCGGACCTGCTTTAT 416

QY 301 GTGCTATTGCCCTTCTCATCATCACCTACTCATCTAGATGACAGGTGGCCCTTCGGGGAG 360
Db 417 GTGCTATTGCCCTTCTCATCATCACCTACTCATCTAGATGACAGGTGGCCCTTCGGGGAG 476

QY 361 CTGCTCTGCAAGCTGGTGACCTTCTCTGTCTTATATCAACCTTTTAGGCGCAGCATCTCTGCTG 420
```

```
Db 477 CTGCTCTGCAAGCTGGTGACATCTCTGTCTTATATCAACCTTTACGGCAGCATCTCTGCTG 536
|||||
QY 421 CTGACCTGCATCTCTGTGCACAGTTCTAGGTGTGTCACCCACCTGTGTTCGTGTCGCC 480
|||||
Db 537 CTGACCTGCATCTCTGTGCACAGTTCTAGGTGTGTCACCCACCTGTGTTCGTGTCGCC 596
|||||
QY 481 TACCGGACCCGAGGAGCATGCTGTGGCTGGCCACAGCACCACTGGGCGCTGGTGGTCTC 540
|||||
Db 597 TACCGGACCCGAGGAGCATGCTGTGGCTGGCCACAGCACCACTGGGCGCTGGTGGTCTC 656
|||||
QY 541 CAGCTGTGTCGCCACACTGGCTTCTCCACACGAGCTACATCAATGGCCACAGATGATCTGG 600
|||||
Db 657 CAGCTGTGTCGCCACACTGGCTTCTCCACACGAGCTACATCAATGGCCACAGATGATCTGG 716
|||||
QY 601 TATGACATGACAGCCAAAGAGAAATTTTATCGGCTTTTGTGCTACGGCATAGTTCTTGACA 660
|||||
Db 717 TATGACATGACAGCCAAAGAGAAATTTTATCGGCTTTTGTGCTACGGCATAGTTCTTGACA 776
|||||
QY 661 TTGCTGTGCTTCTTCTCCCTCTGCTGCTATTTGGTGTGCTATTTCACTGATGCTCAGGAG 720
|||||
Db 777 TTGCTGTGCTTCTTCTCCCTCTGCTGCTATTTGGTGTGCTATTTCACTGATGCTCAGGAG 836
|||||
QY 721 COTGATCAAGCCAGAGGAGAACTCTCATGAGACAGGCAACACAGCCCGAGCCAGGTCCAT 780
|||||
Db 837 COTGATCAAGCCAGAGGAGAACTCTCATGAGACAGGCAACACAGCCCGAGCCAGGTCCAT 896
|||||
QY 781 CCGGACCATCTTACTGTGTGTGCTCTTCACTCTGTTTGTGCTTCCATATATCAC 840
|||||
Db 897 CCGGACCATCTTACTGTGTGTGCTCTTCACTCTGTTTGTGCTTCCATATATCAC 956
|||||
QY 841 TCGCTCTCTTCTACCTCACCATCTGCTTCTGCTTCTCAGGACCTGCCAGCTCTTTGATGGC 900
|||||
Db 957 TCGCTCTCTTCTACCTCACCATCTGCTTCTGCTTCTCAGGACCTGCCAGCTCTTTGATGGC 1016
|||||
QY 901 AGCCAGTGTGCTTACAAAGATATGAGGCTCTCTGCTGTGTGAGTGTGAGTGTGCTTCAACCC 960
|||||
Db 1017 AGCCAGTGTGCTTACAAAGATATGAGGCTCTCTGCTGTGTGAGTGTGAGTGTGCTTCAACCC 1076
|||||
QY 961 AGTCTGTGCTTCTTCTTCAAGGGGGGCAAAATAGAGTGTGAGTGTGCTTCAAGAACTGAGG 1020
|||||
Db 1077 AGTCTGTGCTTCTTCTTCAAGGGGGGCAAAATAGAGTGTGAGTGTGCTTCAAGAACTGAGG 1136
|||||
QY 1021 CAGAACAGTTGGTGAGCATCCAGCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCT 1080
|||||
Db 1137 CAGAACAGTTGGTGAGCATCCAGCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCT 1196
|||||
QY 1081 GG 1082
|||
Db 1197 GG 1198
|||||

RESULT 5
AB065670 1334 bp DNA linear PRI 23-JUL-2002
LOCUS
DEFINITION Homo sapiens gene for seven transmembrane helix receptor,
isolate:CBRC/TM_233.
ACCESSION AB065670
VERSION AB065670.1 GI:21928614
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S.,
Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
TITLE Genome-wide discovery and analysis of human seven transmembrane
helix receptor genes
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1334)
AUTHORS Suwa,M.
TITLE Direct Submission
```

JOURNAL Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/, Tel:81-3-3599-8080, Fax:81-3-3599-8081)

COMMENT This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of gene finding (GeneDecoder), sequence search, motif-domain assignment and transmembrane helix prediction.

And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAST), University of Tokyo].

FEATURES

source
1..1334
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="CBRC7M 233"
/db_xref="taxon:9606"
/chromosome="3"
misc_feature
201..1134
/product="seven transmembrane helix receptor"
/evidence="not_experimental"

ORIGIN

Query Match 100.0%; Score 1082; DB 9; Length 1334;
Best Local Similarity 100.0%; Pred. No. 6.8e-267; Indels 0; Gaps 0;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCCTATTTTGGCTTCCAGGGGAAGCAGAGCGGAGCCGTCGTGGAGCTCTG 60
DB 117 ATGCTGTCCTATTTTGGCTTCCAGGGGAAGCAGAGCGGAGCCGTCGTGGAGCTCTG 176
QY 61 CTCCTGAGGAGGCTCCGGGACATGGAGAGTGGACATGAATACATCAGAGGAACAA 120
DB 177 CTCCTGAGGAGGCTCCGGGACATGGAGAGTGGACATGAATACATCAGAGGAACAA 236
QY 121 GGTCTCTGCCAGTCTTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGTATC 180
DB 237 GGTCTCTGCCAGTCTTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGTATC 296
QY 181 ATCTTTATCTAGGGCTGGCACTAAATGGCACTGTCTTTGGCACTCTCTGGGCGCAAAAC 240
DB 297 ATCTTTATCTAGGGCTGGCACTAAATGGCACTGTCTTTGGCACTCTCTGGGCGCAAAAC 356
QY 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGAACCTGATGTGGCGGACCTGCTTTAT 300
DB 357 AAGCGTGGAGCTGTGCCACCACTATCTGTGAACCTGATGTGGCGGACCTGCTTTAT 416
QY 301 GTGCTATTTGCCCTTCCTCATCATCACTACTCATAGATGACAGGTGGCCCTTCGGGGAG 360
DB 417 GTGCTATTTGCCCTTCCTCATCATCACTACTCATAGATGACAGGTGGCCCTTCGGGGAG 476
QY 361 CTGCTCTGCAAGCTGTGGCACTCTCTCTCTATATCACTTACCTTTACGCGACATCTGCTG 420
DB 477 CTGCTCTGCAAGCTGTGGCACTCTCTCTCTATATCACTTACGCGACATCTGCTGCTG 536
QY 421 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTGTGCTGCTGCC 480
DB 537 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTGTGCTGCTGCC 596
QY 481 TACCGGACCGGACGCTGCTGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 597 TACCGGACCGGACGCTGCTGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 656
QY 541 CAGTGTGTGCCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 657 CAGTGTGTGCCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 716
QY 601 TATGACATGACCAAGAGAAATTTTGAATGGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 660

DB 717 TATGACATGACCAAGAGAAATTTTGTATCGGCTTTTGTCTACGGCATATGCTGACA 776
QY 661 TTGCTGTGCTTCTTTTCCCTCTTGGTCAATTTTGGTGTGCTATTCATGATGGTCAGAG 720
DB 777 TTGCTGTGCTTCTTTTCCCTCTTGGTCAATTTTGGTGTGCTATTCATGATGGTCAGAG 836
QY 721 COTGATCAAGCAGAGGAGAACCTCATGAGCAGGCAACACAGCCCGAGCCAGTCCAT 780
DB 837 COTGATCAAGCAGAGGAGAACCTCATGAGCAGGCAACACAGCCCGAGCCAGTCCAT 896
QY 781 CCGGACCATCTCTACGTGTGTGGCTCTTCAACCTCTGTGTTTGTGCTGCTTCCATATCAC 840
DB 897 CCGGACCATCTCTACGTGTGTGGCTCTTCAACCTCTGTGTTTGTGCTGCTTCCATATCAC 956
QY 841 TCGCTCTCTTACCTCAACATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
DB 957 TCGCTCTCTTACCTCAACATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1016
QY 901 AGCCAGTGTGGCTACAAGATATGAGGCTCTGCTGTGAGTGTGAGCAGCTGCTCAACCC 960
DB 1017 AGCCAGTGTGGCTACAAGATATGAGGCTCTGCTGTGAGTGTGAGCAGCTGCTCAACCC 1076
QY 961 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTGAGG 1020
DB 1077 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTGAGG 1136
QY 1021 CAGAACTAGTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGTTGAAACAGATCT 1080
DB 1137 CAGAACTAGTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGTTGAAACAGATCT 1196
QY 1081 GG 1082
DB 1197 GG 1198

RESULT 6
LOCUS AX365511 3143 bp DNA linear PAT 21-FEB-2002
DEFINITION Sequence 1 from Patent WO0185764.
ACCESSION AX365511
VERSION AX365511.1 GI:18873666
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Ramakrishnan, S.
TITLE Regulation of human p2y-like g protein-coupled receptor
JOURNAL Patent: WO 0185764-A 1 15-NOV-2001;
Bayer Aktiengesellschaft (DE)
LOCATION/Qualifiers
1..3143
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
520..>1374
/note="unnamed protein product"

FEATURES

source
1..3143
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
520..>1374
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD24312.1"
/db_xref="GI:18873667"

CDS

ORIGIN

Query Match 100.0%; Score 1082; DB 6; Length 3143;
Best Local Similarity 100.0%; Pred. No. 7.3e-267; Indels 0; Gaps 0;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCCATTTTCTTCTTCTCAGGGAGACAGAGCGGAGCGCTGCTGGAGCTCTG 60
 Db 436 ATGCTGTCCATTTTCTTCTTCTCAGGGAGACAGAGCGGAGCGCTGCTGGAGCTCTG 495
 QY 61 CTCCTGGAGGAGCTCTCCGGGACATGGAGAGGTGGACATGATATACATACAGGACAA 120
 Db 496 CTCCTGGAGGAGCTCTCCGGGACATGGAGAGGTGGACATGATATACATACAGGACAA 555
 QY 121 GGTCTCTGCCAGTTCTCAGAGAGTACAGAGAGTCTACCTCTCTCCCTGGGCTACAGTATC 180
 Db 556 GGTCTCTGCCAGTTCTCAGAGAGTACAGAGAGTCTACCTCTCTCCCTGGGCTACAGTATC 615
 QY 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTCTCTTGTGGCACTCTCTGGGGCCAAACC 240
 Db 616 ATCTTTATCTAGGCTGCCACTAAATGGCACTCTCTTGTGGCACTCTCTGGGGCCAAACC 675
 QY 241 AAGCGCTGGAGCTGGCCACCACTATCTGGTGAACCTGATGTGGCGGACCTGCTTTAT 300
 Db 676 AAGCGCTGGAGCTGGCCACCACTATCTGGTGAACCTGATGTGGCGGACCTGCTTTAT 735
 QY 301 GTGCTATTGGCCCTTCTCATCATCCTACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 360
 Db 736 GTGCTATTGGCCCTTCTCATCATCCTACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 795
 QY 361 CTGCTCTGCAAGCTGGTGCCTCTCTCTTCTATATCAACCTTTACGGCAGCATCTGCTG 420
 Db 796 CTGCTCTGCAAGCTGGTGCCTCTCTCTTCTATATCAACCTTTACGGCAGCATCTGCTG 855
 QY 421 CTGACCTGCATCTCTGTGCACAGTCTCTAGGTGTGTGCACCACTGTGTGTGCTGCC 480
 Db 856 CTGACCTGCATCTCTGTGCACAGTCTCTAGGTGTGTGCACCACTGTGTGTGCTGCC 915
 QY 481 TACCGGACCCGAGCATCTGCTGGTGGGACCAAGCAACCTGGGCGCTGGTGGTCTC 540
 Db 916 TACCGGACCCGAGCATCTGCTGGTGGGACCAAGCAACCTGGGCGCTGGTGGTCTC 975
 QY 541 CAGTGTCTGCCACATCTGGCTCTTCCACACGAGCTACATCAATGACGACATGATCTGG 600
 Db 976 CAGTGTCTGCCACATCTGGCTCTTCCACACGAGCTACATCAATGACGACATGATCTGG 1035
 QY 601 TATGACATGACGACGACGAGATTTTGTGCGCTTTTGTGCTACGGCATAGTTCTGACA 660
 Db 1036 TATGACATGACGACGACGAGATTTTGTGCGCTTTTGTGCTACGGCATAGTTCTGACA 1095
 QY 661 TTGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
 Db 1096 TTGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1155
 QY 721 CCTGATCAAGGAGGAGACCTCTGAGGACAGGCAACAGCCCGGAGGCTCCAT 780
 Db 1156 CCTGATCAAGGAGGAGACCTCTGAGGACAGGCAACAGCCCGGAGGCTCCAT 1215
 QY 781 CCGGACCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
 Db 1216 CCGGACCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1275
 QY 841 TCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
 Db 1276 TCGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1335
 QY 901 AGCAGTGTGGCTTACAGATATGAGGCGCTCTGGTGTGTGTGTGTGTGTGTGTGTGTGT 960
 Db 1336 AGCAGTGTGGCTTACAGATATGAGGCGCTCTGGTGTGTGTGTGTGTGTGTGTGTGTGT 1395
 QY 961 AGTCTGTACTTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGGTCTCTCCAGAACTGAGG 1020
 Db 1396 AGTCTGTACTTCTTCTTCAAGGGGGGCAAAATAGAGTCAAGGTCTCTCCAGAACTGAGG 1455
 QY 1021 CAGACAGTGTGGTGTGAGCTTCCAGCTGGGAGGAGAGATGCCAGGGTTGAACATCT 1080
 Db 1456 CAGACAGTGTGGTGTGAGCTTCCAGCTGGGAGGAGAGATGCCAGGGTTGAACATCT 1515
 QY 1081 GG 1082

Db 1516 GG 1517

RESULT 7
 AC026960/c
 LOCUS
 DEFINITION
 AC026960
 Homo sapiens chromosome 3 clone RP11-170K4 map 3, WORKING DRAFT
 SEQUENCE, 15 unordered pieces.
 AC026960.2 GI:7547202
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 3, clone RP11-170K4
 Unpublished
 2 (bases 1 to 164502)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,E.,
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kahn,L., Karatas,A.,
 Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marguis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Olivaz,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigglio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 13, 2000 this sequence version replaced gi:7328839.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L8491
 Center clone name: 170_K_4
 ----- Summary Statistics
 Sequencing vector: M13; M77815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 156927 bases at least Q40
 Consensus quality: 160376 bases at least Q30
 Consensus quality: 161800 bases at least Q20
 Insert size: 170000; agarose-fp
 Insert size: 163102; sum-of-ctnigs
 Quality coverage: 4.4 in Q20 bases; agarose-fp
 Quality coverage: 4.5 in Q20 bases; sum-of-ctnigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1 2917: contig of 2917 bp in length
2918 3017: gap of 100 bp
3018 3018: contig of 3293 bp in length
6311 6410: gap of 100 bp
6411 9447: contig of 3037 bp in length
9448 9548: gap of 100 bp
9548 14470: contig of 4923 bp in length
14471 14570: gap of 100 bp
14571 20528: contig of 5958 bp in length
20529 20628: gap of 100 bp
20629 20629: contig of 4039 bp in length
20630 24568: gap of 100 bp
24568 24568: contig of 4600 bp in length
24569 29367: gap of 100 bp
29368 29468: contig of 6770 bp in length
29468 36237: gap of 100 bp
36238 36338: contig of 10040 bp in length
36338 46477: gap of 100 bp
46478 59361: contig of 12884 bp in length
59362 59461: gap of 100 bp
59462 75221: contig of 15760 bp in length
75222 75321: gap of 100 bp
75322 93371: contig of 18050 bp in length
93372 93471: gap of 100 bp
93472 117463: contig of 23992 bp in length
117464 117563: gap of 100 bp
117564 137715: contig of 20152 bp in length
137716 137716: gap of 100 bp
137816 164502: contig of 26687 bp in length.
```

FEATURES

source

```
1. 164502
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="3"
/map="3"
/clone="RP11-170K4"
/clone_lib="RP11-11 Human Male BAC"
1. 2517
/notes="assembly_fragment"
3018. 6310
/notes="assembly_fragment"
6411. 9447
/notes="assembly_fragment"
9548. 14470
/notes="assembly_fragment"
14571. 20528
/notes="assembly_fragment"
20629. 24667
/notes="assembly_fragment"
clone_end:SP6
vector_side:right
24768. 29367
/notes="assembly_fragment"
29468. 36237
/notes="assembly_fragment"
36338. 46377
/notes="assembly_fragment"
46478. 59361
/notes="assembly_fragment"
59462. 75221
/notes="assembly_fragment"
75322. 93371
/notes="assembly_fragment"
93472. 117463
/notes="assembly_fragment"
117564. 137715
/notes="assembly_fragment"
clone_end:T7
```

```
misc_feature      vector_side:left"
137816..164502
/notes="assembly_fragment"

ORIGIN
Query Match      100.0%; Score 1082; DB 2; Length 164502;
Best Local Similarity 100.0%; Pred. No. 1e-266;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCCATTTTCTTCCAGGGGAGAGAGAGCGGGAGCCCTCGTGAGACTCTG 60
Db 123728 ATGCTGTCCATTTTCTTCCAGGGGAGAGAGAGCGGGAGCCCTCGTGAGACTCTG 123669
QY 61 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCACAGGAACAA 120
Db 123668 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCACAGGAACAA 123609
QY 121 GGTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 180
Db 123608 GGTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 123549
QY 181 ATCTTTATCCTAGGGCTGCCACTAAATGGCACTGTCTTTGTGGCACTCTCTGGGGCAAAACC 240
Db 123548 ATCTTTATCCTAGGGCTGCCACTAAATGGCACTGTCTTTGTGGCACTCTCTGGGGCAAAACC 123489
QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCGGACCTGCTTTAT 300
Db 123488 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCGGACCTGCTTTAT 123429
QY 301 GTGCTATTGGCCCTTCTCATCATCCTACTCATCTAGATGACAGGTGGCCCTTCGGGGGAG 360
Db 123428 GTGCTATTGGCCCTTCTCATCATCCTACTCATCTAGATGACAGGTGGCCCTTCGGGGGAG 123369
QY 361 CTGCTGTCAAGCTGGTGCACTTCTGTCTATATCAACCTTTACGGGAGCATCTGCTGT 420
Db 123368 CTGCTGTCAAGCTGGTGCACTTCTGTCTATATCAACCTTTACGGGAGCATCTGCTGT 123309
QY 421 CTGACCTGCACTCTGTGCACTGCTCTAGTGTGTGCCACCACTGTGTTCGCTGCC 480
Db 123308 CTGACCTGCACTCTGTGCACTGCTCTAGTGTGTGCCACCACTGTGTTCGCTGCC 123249
QY 481 TACCGGACCCGAGGCATGCTGCTGGGCAACAGCAACCTCTGGGCCCTTGGTGGTCTC 540
Db 123248 TACCGGACCCGAGGCATGCTGCTGGGCAACAGCAACCTCTGGGCCCTTGGTGGTCTC 123189
QY 541 CAGCTGTGCCCACTGCTGGCCCTTCTCCACAGCACTACATCAATGCCAGATCATCTGG 600
Db 123188 CAGCTGTGCCCACTGCTGGCCCTTCTCCACAGCACTACATCAATGCCAGATCATCTGG 123129
QY 601 TATGACATGACAGCAAGAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGACA 660
Db 123128 TATGACATGACAGCAAGAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGACA 123069
QY 661 TTGCTGGCTTTCTTTCCCTCCTTGGTCAATTTTGGTGTGCTATTTCATGTGTCAGGAG 720
Db 123068 TTGCTGGCTTTCTTTCCCTCCTTGGTCAATTTTGGTGTGCTATTTCATGTGTCAGGAG 123009
QY 721 CCTGATCAAGCCAGAGAGAACCTCATGAGGAGAGCAACACAGCCGAGCCAGGTCCAT 780
Db 123008 CCTGATCAAGCCAGAGAGAACCTCATGAGGAGAGCAACACAGCCGAGCCAGGTCCAT 122949
QY 781 CCGGACCATCTCTACTGTGTGTGGCCCTCTTCACTCTGTGTTTGTGGCCCTTCCATATCAC 840
Db 122948 CCGGACCATCTCTACTGTGTGTGGCCCTCTTCACTCTGTGTTTGTGGCCCTTCCATATCAC 122889
QY 841 TGGCTCTTCTACCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGCG 900
Db 122888 TGGCTCTTCTACCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGCG 122829
QY 901 AGCCAGTGTGGCCTACAAGATATGGAGGCTCTCTGTGTGAGTGTGAGCAGCTCCCTCAACC 960
Db 122828 AGCCAGTGTGGCCTACAAGATATGGAGGCTCTCTGTGTGAGTGTGAGCAGCTCCCTCAACC 122769
```

```

QY 961 AGTCCTGTACTTTCTTTCAAGGGGGGCAAAAATAGAGTCAGGCTCTCCAGAACTGAGG 1020
DB 122768 AGTCCTGTACTTTCTTTCAAGGGGGGCAAAAATAGAGTCAGGCTCTCCAGAACTGAGG 122709
QY 1021 CAGAACAGTTGGTGGAGTCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT 1080
DB 122708 CAGAACAGTTGGTGGAGTCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT 122649
QY 1081 GG 1082
DB 122648 GG 122647

RESULT 8
LOCUS AC092999 167084 bp DNA linear PRI 30-JUL-2002
DEFINITION Homo sapiens 3 BAC RP11-170K4 (Roswell Park Cancer Institute Human
ACCESSION AC092999 BAC Library) complete sequence.
VERSION AC092999.3 GI:19033407
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 167084)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oman,F.R., Allen,C.,
Alsbrooks,S.B., Amarantunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burckett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Edwards,C.C., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Guaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homsai,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
Mei,G., Metsker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nickerson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenwo,S., Ogut,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Face,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pockens,R., Primus,E., Pu,L.B., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E.,
Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 167084)
Worley,K.C.
Direct Submission
Submitted (09-AUG-2001) Human Genome Sequencing Center, Department

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

```

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 167084)
 Worley,K.C.
 Direct Submission
 Submitted (01-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 167084)
 Worley,K.C.
 Direct Submission
 Submitted (29-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 5 (bases 1 to 167084)
 Worley,K.C.
 Direct Submission
 Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On Mar 1, 2002 this sequence version replaced gi:18139250.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
 STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
 Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
 Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.
 Location/Qualifiers
 1..167084
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-170K4"
 complement(1..2005)
 /note="overlaps bases 1..2005 of clone AC092997"
 /function="clone overlap"
 189..376
 /rpt_family="MIR"
 complement(997..1123)
 /rpt_family="L2"

misc_feature
 repeat_region
 repeat_region

repeat_region	1999..2058	/rpt_family="L2"	1	ATGCTGTCATTTTGTCTTCCCTTCAGGGGAGCAGAGCGGAGCCCTCGTGGAGCTCTG	60
repeat_region	2096..2404	/rpt_family="AluSx"	160921	ATGCTGTCATTTTGTCTTCCCTTCAGGGGAGCAGAGCGGAGCCCTCGTGGAGCTCTG	160980
repeat_region	2422..2529	/rpt_family="L2"	61	CTCCTGGAGGGAGCCTCCCGGACATGGAGAGGTGGACATGATATCATCATCAGAACAA	120
repeat_region	2975..3004	/rpt_family="AT_rich"	160981	CTCCTGGAGGGAGCCTCCCGGACATGGAGAGGTGGACATGATATCATCATCAGAACAA	161040
repeat_region	3134..3163	/rpt_family="AT_rich"	121	GGTCTCTGCCAGTTCTCTCAGAGAGTCAAGCAAGTCTACTCTCTCCCTGGCCCTACAGTATC	180
repeat_region	4270..4290	/rpt_family="AT_rich"	161041	GGTCTCTGCCAGTTCTCTCAGAGAGTCAAGCAAGTCTACTCTCTCCCTGGCCCTACAGTATC	161100
repeat_region	complement(4421..4550)	/rpt_family="MIR"	181	ATCTTTATCTAGGGCTGCCAATAATGGGACCTCTCTTGTGGACCTCTCTGGGGCAAAACC	240
repeat_region	5251..5642	/rpt_family="MLT1K"	161101	ATCTTTATCTAGGGCTGCCAATAATGGGACCTCTCTTGTGGACCTCTCTGGGGCAAAACC	161160
repeat_region	complement(5812..6366)	/rpt_family="MER34B"	241	AAGCGCTGGAGCTGTGCCACCACCTATCTGTGTGAACCTGTGTGGCGGACCTCTCTTTAT	300
repeat_region	6909..6941	/rpt_family="AT_rich"	161161	AAGCGCTGGAGCTGTGCCACCACCTATCTGTGTGAACCTGTGTGGCGGACCTCTCTTTAT	161220
repeat_region	7595..7717	/rpt_family="MERSA"	301	GTGCTATTGCCCTTCTCATCATCACCTACTACTAGATGACAGGTGGCCCTTCGGGGAG	360
repeat_region	complement(7788..8198)	/rpt_family="MLT1I"	161221	GTGCTATTGCCCTTCTCATCATCACCTACTACTAGATGACAGGTGGCCCTTCGGGGAG	161280
repeat_region	complement(8536..8591)	/rpt_family="L2"	361	CTGCTCTGCAAGCTGTGCACTTCTCTTCTATATCAACCTTTTACGGCAGCATCTCTG	420
repeat_region	9634..9936	/rpt_family="AluSx"	161281	CTGCTCTGCAAGCTGTGCACTTCTCTTCTATATCAACCTTTTACGGCAGCATCTCTG	161340
repeat_region	complement(10766..10814)	/rpt_family="MIR"	421	CTGACCTGCACTCTGTGTGCCACCTCTCTAGGTGTGTGCCACCACTGTGTCTGCTGCC	480
repeat_region	complement(10848..10964)	/rpt_family="MIR"	161341	CTGACCTGCACTCTGTGTGCCACCTCTCTAGGTGTGTGCCACCACTGTGTCTGCTGCC	161400
repeat_region	17399..17431	/rpt_family="(TTAAA)n"	481	TACGGACCCGACGACATGCCCTGTGGCCACGACACCTGGGCGCTTGGTGGCTCTC	540
repeat_region	complement(17432..17546)	/rpt_family="MIR"	161401	TACGGACCCGACGACATGCCCTGTGGCCACGACACCTGGGCGCTTGGTGGCTCTC	161460
repeat_region	17637..18073	/rpt_family="Charlieb"	541	CAGCTGTGCCACACACTGGCCCTTCTCCACACGAGTACATCAATGGCCAGATGATCTGG	600
repeat_region	18278..18313	/rpt_family="(TG)n"	161461	CAGCTGTGCCACACACTGGCCCTTCTCCACACGAGTACATCAATGGCCAGATGATCTGG	161520
repeat_region	21428..21841	/rpt_family="MSTD"	601	TATGACATGACCGACCAAGAGATTTTGTGGCTTTTGTGGCTTACGAGTAGTTCTGACA	660
repeat_region	complement(21847..21926)	/rpt_family="LTR33A"	161521	TATGACATGACCGACCAAGAGATTTTGTGGCTTTTGTGGCTTACGAGTAGTTCTGACA	161580
repeat_region	22387..22631	/rpt_family="MLT1G1"	661	TTGCTGGCTTTCTTCCCTCCTTGGTCTATTTGGTGTGCTATTCACCTGATGTCAGGAG	720
repeat_region	complement(22768..23212)	/rpt_family="Tigger3(Golem)"	161581	TTGCTGGCTTTCTTCCCTCCTTGGTCTATTTGGTGTGCTATTCACCTGATGTCAGGAG	161640
repeat_region	complement(23222..23279)	/rpt_family="Tigger3(Golem)"	721	CCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACAACAGCCCGAGCCAGGTCCAT	780
repeat_region	23311..23437	/rpt_family="MLT1E2"	161641	CCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACAACAGCCCGAGCCAGGTCCAT	161700
repeat_region	24105..24225	/rpt_family="MIR"	781	CGGACCATCTACTGGTGTGGCTTCTCACCTCTGTGTTTGGCTTCCATATCAC	840
repeat_region	24226..24517	/rpt_family="MER33"	161701	CGGACCATCTACTGGTGTGGCTTCTCACCTCTGTGTTTGGCTTCCATATCAC	161760
repeat_region	24518..24629	/rpt_family="MIR"	841	TCGCTCTCTTCTACCTCACCATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGGC	900
repeat_region	complement(24639..24724)	/rpt_family="L2"	161761	TCGCTCTCTTCTACCTCACCATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGGC	161820
repeat_region	25420..25994	/rpt_family="L2"	901	AGCCAGTGTGGCTTACAGATATGAGGCTCTCTGGTGTGAGTGTGAGAGCTTCCCTCAACCC	960
repeat_region	26024..26059	/rpt_family="(TG)n"	161821	AGCCAGTGTGGCTTACAGATATGAGGCTCTCTGGTGTGAGTGTGAGAGCTTCCCTCAACCC	161880
repeat_region	26565..27405	/rpt_family="L1MB8"	961	AGTCTGTACTTTCTTTTCAAGGGGGCAAAATAGAGTCAGGCTCTTCCAGAAACTGAGG	1020
repeat_region	complement(27789..27855)	/rpt_family="L2"	161881	AGTCTGTACTTTCTTTTCAAGGGGGCAAAATAGAGTCAGGCTCTTCCAGAAACTGAGG	161940
repeat_region	28628..28778		1021	CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAAACAGATCT	1080
repeat_region			161941	CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAAACAGATCT	162000

Query Match 100.0%; Score 1082; DB 9; Length 167084;
 Best Local Similarity 100.0%; Pred. No. 1e-256;
 Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
Query Match 99.0%; Score 1071; DB 2; Length 163958;
Best Local Similarity 99.9%; Pred. No. 6 8e-264; Indels 1; Gaps 1;
Matches 1082; Conservative 0; Mismatches 0;

QY 1 ATGCTGTCCATTTTGTCTTCCAGGGAAGCAGAGCGGAGCGTGTGGAGCTCTG 60
Db 98103 ATGCTGTCCATTTTGTCTTCCAGGGAAGCAGAGCGGAGCGTGTGGAGCTCTG 98162
QY 61 CTCTCTGAGGAGCTCCCGGACATGAGAGGTGACATGATACATACAGAGCAAA 120
Db 98163 CTCTCTGAGGAGCTCCCGGACATGAGAGGTGACATGATACATACAGAGCAAA 98222
QY 121 GGTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGCCCTACAGTATC 180
Db 98223 GGTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGCCCTACAGTATC 98282
QY 181 ATCTTTATCTAGGCTGCCACTAAATGGACATGCTTGTGGACATCTCTGGGCGCAACC 240
Db 98283 ATCTTTATCTAGGCTGCCACTAAATGGACATGCTTGTGGACATCTCTGGGCGCAACC 98342
QY 241 AAGCGCTGAGCTGTGCCACCCTATCTGGTGAACCTGATGTGGCCGACCTGCTTAT 300
Db 98343 AAGCGCTGAGCTGTGCCACCCTATCTGGTGAACCTGATGTGGCCGACCTGCTTAT 98402
QY 301 GTGCTATGCTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTGGGGAG 360
Db 98403 GTGCTATGCTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTGGGGAG 98462
QY 361 CTGCTCTGAGCTGTGGACTCTCTGTTCTATCAACTTTACGGCAGCATCTGCTG 420
Db 98463 CTGCTCTGAGCTGTGGACTCTCTGTTCTATCAACTTTACGGCAGCATCTGCTG 98522
QY 421 CTGACCTGATCTCTGTGACCAAGTCTCTAGGTGTGTGCCACCCTGTTGTGCTGCC 480
Db 98523 CTGACCTGATCTCTGTGACCAAGTCTCTAGGTGTGTGCCACCCTGTTGTGCTGCC 98582
QY 481 TACGGACCGGAGGATCTGCTGGTGGGACCAAGCAGCAGTCTGGGCGCTGGTCTTC 540
Db 98583 TACGGACCGGAGGATCTGCTGGTGGGACCAAGCAGCAGTCTGGGCGCTGGTCTTC 98642
QY 541 CAGCTGTGCCACACTGTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 600
Db 98643 CAGCTGTGCCACACTGTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 98702
QY 601 TATGACATGACAGCAAGAGAAATTTGTATGCGCTTTTGTGCGCATAGTTCTGACA 660
Db 98703 TATGACATGACAGCAAGAGAAATTTGTATGCGCTTTTGTGCGCATAGTTCTGACA 98762
QY 661 TTGCTGGCTTTCTTTCCCTCTCTTGTGCTATTTTGTGCTATTTCACTGATGTCAGGAG 720
Db 98763 TTGCTGGCTTTCTTTCCCTCTCTTGTGCTATTTTGTGCTATTTCACTGATGTCAGGAG 98822
QY 721 CCTGATCAAGCCAGAGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCCAGTCCAT 780
Db 98823 CCTGATCAAGCCAGAGAGAACCTCATGAGGACAGGCAACACAGCCCGAGGTCAT 98882
QY 781 CCGGACCATCTACTGCTGTGGCTCTTACCCCTGTTTGTGGCTTTCATATCAC 840
Db 98883 CCGGACCATCTACTGCTGTGGCTCTTACCCCTGTTTGTGGCTTTCATATCAC 98942
QY 841 TCGCTCTCTTACTCATCATCTGCTTTCTGCTTTTCTCAGGACTGCGCAGCTCTTATGCG 900
Db 98943 TCGCTCTCTTACTCATCATCTGCTTTCTGCTTTTCTCAGGACTGCGCAGCTCTTATGCG 99002
QY 901 AG-CCAGTGTGGCTTCAAGATATGAGGCTCTCTGTTAGTGTGAGCAGCTGCTCAACC 959
Db 99003 AGCCCACTGTGGCTTCAAGATATGAGGCTCTCTGTTAGTGTGAGCAGCTGCTCAACC 99062
QY 960 CAGTCTGTACTTTCTTTCAAGGGGGCAAAATAGAGTCAAGTCAAGTCTCTCCAGAACTGAG 1019
Db 99063 CAGTCTGTACTTTCTTTCAAGGGGGCAAAATAGAGTCAAGTCAAGTCTCTCCAGAACTGAG 99122
QY 1020 GCAGAAACAAGTGGTGAGCATCCAGCTGGGAGAGAGATGCCAGGTTGAACAGATC 1079
```

```
Db 99123 GCAGAAACAAGTGGTGAGCATCCAGCTGGGAGAGAGATGCCAGGTTGAACAGATC 99182
QY 1080 TGG 1082
Db 99183 TGG 99185

RESULT 10
AX921825 AX921825 1108 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 165 from Patent WO02068649.
DEFINITION AX921825
ACCESSION AX921825
VERSION AX921825.1 GI:40215364
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Patent: WO 02068649-A 165 06-SEP-2002;
JOURNAL Curagen Corporation (US)
FEATURES
source
1..1108
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 97.9%; Score 1059; DB 6; Length 1108;
Best Local Similarity 99.8%; Pred. No. 5.5e-261;
Matches 1081; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
QY 1 ATGCTGTCCATTTTGTCTTCCAGGGAAGCAGAGCGGAGCGTGTGGAGCTCTG 60
Db 12 ATGCTGTCCATTTTGTCTTCCAGGGAAGCAGAGCGGAGCGTGTGGAGCTCTG 71
QY 61 CTCCTGAGGAGGAGCTCCCGGACATGAGAGAGTGTGACATGATACATACAGAGCAAA 120
Db 72 CTCCTGAGGAGGAGCTCCCGGACATGAGAGAGTGTGACATGATACATACAGAGCAAA 131
QY 121 GGTCTCTGCCAGTTCTCAGAGAACTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
Db 132 GGTCTCTGCCAGTTCTCAGAGAACTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 191
QY 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTCTTGTGGCACTCTCTGGGCGCAACC 240
Db 192 ATCTTTATCTAGGCTGCCACTAAATGGCACTCTTGTGGCACTCTCTGGGCGCAACC 251
QY 241 AAGCGCTGAGCTGTGCCACCCTATCTTGTGAACCTGATGTGGCGGAGCTCTTTAT 300
Db 252 AAGCGCTGAGCTGTGCCACCCTATCTTGTGAACCTGATGTGGCGGAGCTCTTTAT 311
QY 301 GTGCTATTGCTCTCTCATCATCACCTTACTACTAGATGACAGGTGGCCCTTGGGGAG 360
Db 312 GTGCTATTGCTCTCTCATCATCACCTTACTACTAGATGACAGGTGGCCCTTGGGGAG 371
QY 361 CTGCTCTGCAAGCTGTGGTGCATCTCTGTTCTATATCAACCTTTACGGCAGCATCTGCTG 420
Db 372 CTGCTCTGCAAGCTGTGGTGCATCTCTGTTCTATATCAACCTTTACGGCAGCATCTGCTG 431
QY 421 CTGACCTCATCTCTGTGACCAAGTCTCTAGGTGTGTGGCAACCTGTTCTGCTGCC 480
Db 432 CTGACCTCATCTCTGTGACCAAGTCTCTAGGTGTGTGGCAACCTGTTCTGCTGCC 491
QY 481 TACCGGACCGGAGGAGTGTGCTGGTGGGCAACAGCAACCCTGCGGCTGTGGTGTCTC 540
Db 492 TACCGGACCGGAGGAGTGTGCTGGTGGGCAACAGCAACCCTGCGGCTGTGGTGTCTC 551
QY 541 CAGCTGTGCTGCCACACTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 600
Db 552 CAGCTGTGCTGCCACACTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 611
```

QY 601 TATGACATGACCCAGAGAAATTTTGATCGGCTTTTCCCTACCGCATAGTCTGACA 660
 Db 612 TATGACATGACCCAGAGAAATTTTGATCGGCTTTTCCCTACCGCATAGTCTGACA 671
 QY 661 TTGCTCGCTTTCTTTCCCTCGCTGCTCAATTTGGTGTGCTATTCATGATGCTCAGGAG 720
 Db 672 TTGCTCGCTTTCTTTCCCTCGCTGCTCAATTTGGTGTG-TATTCATGATGCTCAGGAG 730
 QY 721 CCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGTCCAT 780
 Db 731 CCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGTCCAT 790
 QY 781 CCGGACCATCTACTGCTGTGTGGCTCTTCAACCTCTGTTTGTGCTTCCATATCAC 840
 Db 791 CCGGACCATCTACTGCTGTGTGGCTCTTCAACCTCTGTTTGTGCTTCCATATCAC 850
 QY 841 TCGCTCTCTTACCTACCATCTGCTTTCTGCTTTCTCAGGACTGCGCAGCTTTGATGGC 900
 Db 851 TCGCTCTCTTACCTACCATCTGCTTTCTGCTTTCTCAGGACTGCGCAGCTTTGATGGC 910
 QY 901 AG-CCAGTGTGGCTACAGATATGAGAGGCTCTGTGAGTGTGAGAGTGTGCTCAACC 959
 Db 911 AGCCAGTGTGGCTACAGATATGAGAGGCTCTGTGAGTGTGAGAGTGTGCTCAACC 970
 QY 960 CAGTCTGTACTTTCTTTCAAGGGGGCAAAATAGAGTCAAGTCTCTCCAGAACTGAG 1019
 Db 971 CAGTCTGTACTTTCTTTCAAGGGGGCAAAATAGAGTCAAGTCTCTCCAGAACTGAG 1030
 QY 1020 GCAGAAACAAGTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGTTGAACAGATC 1079
 Db 1031 GCAGAAACAAGTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGTTGAACAGATC 1090
 QY 1080 TGG 1082
 Db 1091 TGG 1093

RESULT 11
 AF411108 1019 bp DNA linear PRI 01-NOV-2001
 LOCUS Homo sapiens GPR79 pseudogene, complete sequence.
 DEFINITION AF411108
 VERSION AF411108.1 GI:16566321
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Lee, D.K., Nguyen, T., Lynch, K.R., Cheng, R., Vanti, W.B., Arkhitko, O.,
 Lewis, T., Evans, J.F., George, S.R. and O'Dowd, B.F.
 TITLE Discovery and mapping of ten novel G protein-coupled receptor genes
 Gene 275 (1), 83-91 (2001)
 MEDLINE 21458557
 PUBMED 11574155
 REFERENCE 2 (bases 1 to 1019)
 AUTHORS Lee, D.K., Nguyen, T., Lynch, K.R., Cheng, R., Vanti, W.B., Arkhitko, O.,
 Lewis, T., Evans, J.F., George, S.R. and O'Dowd, B.F.
 TITLE Direct Submission
 JOURNAL Submitted (17-AUG-2001) Department of Pharmacology, University of
 Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada
 FEATURES
 source 1..1019
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene 1..1019
 /gene="GPR79"
 /pseudo

Query Match 93.1%; Score 1007; DB 9; Length 1019;

Best Local Similarity 99.9%; Pred. No. 1.2e-247;
 Matches 1018; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 ATGCTGTCCATTTTGTCTTCCAGGGGAGAGAGAGCGGAGCCGTCTGTGAGCTCTG 60
 Db 1 ATGCTGTCCATTTTGTCTTCCAGGGGAGAGAGAGCGGAGCCGTCTGTGAGCTCTG 60
 QY 61 CTCCTGAGGAGCCCTCCCGGACATGGAGAGGTGGACATGAATATACACAGAAACA 120
 Db 61 CTCCTGAGGAGCCCTCCCGGACATGGAGAGGTGGACATGAATATACACAGAAACA 120
 QY 121 GGTCTCTCCAGTCTCAGAGAAGTACAAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
 Db 121 GGTCTCTCCAGTCTCAGAGAAGTACAAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
 QY 181 ATCTTTATCTAGGGCTGCCACTAAATGGCACTGTCTTGGGCACTCTCTGGGCAAAACC 240
 Db 181 ATCTTTATCTAGGGCTGCCACTAAATGGCACTGTCTTGGGCACTCTCTGGGCAAAACC 240
 QY 241 AAGGCTCGAGCTGTGCCACCACCTATCTGGTGAACCTGATGCTGGCGGACCTGCTTTAT 300
 Db 241 AAGGCTCGAGCTGTGCCACCACCTATCTGGTGAACCTGATGCTGGCGGACCTGCTTTAT 300
 QY 301 GTGCTATTGCCCTTCTCTCATCATCACTACTAGATGACAGGTGGCCCTTCGGGAG 360
 Db 301 GTGCTATTGCCCTTCTCTCATCATCACTACTAGATGACAGGTGGCCCTTCGGGAG 360
 QY 361 CTGCTCTGCAAGCTGTGCACTTCTCTATATCAACCTTTACGGAGCATCTCTGCTG 420
 Db 361 CTGCTCTGCAAGCTGTGCACTTCTCTATATCAACCTTTACGGAGCATCTCTGCTG 420
 QY 421 CTGACCTGCATCTCTGTGCAACAGTTCCTAGGTGTGTGCCACCCACTGTGTCTGCC 480
 Db 421 CTGACCTGCATCTCTGTGCAACAGTTCCTAGGTGTGTGCCACCCACTGTGTCTGCC 480
 QY 481 TACGGAGCCGAGGATGCTGCTGGGACAGGACACCTGGGCGCTGGTGGTCTC 540
 Db 481 TACGGAGCCGAGGATGCTGCTGGGACAGGACACCTGGGCGCTGGTGGTCTC 540
 QY 541 CAGTGTGCTGCCACACTGGCCCTCTCCACACGAGTACATCAATGGCCAGATCATCTGG 600
 Db 541 CAGTGTGCTGCCACACTGGCCCTCTCCACACGAGTACATCAATGGCCAGATCATCTGG 600
 QY 601 TATGACATGACCCAGAGAAATTTTGATCGGCTTTTGGCTACGGCATAGTCTGACA 660
 Db 601 TATGACATGACCCAGAGAAATTTTGATCGGCTTTTGGCTACGGCATAGTCTGACA 660
 QY 661 TTGCTCGCTTTCTTTCCCTCGCTGCTCAATTTGGTGTGCTATTCATGATGCTCAGGAG 720
 Db 661 TTGCTCGCTTTCTTTCCCTCGCTGCTCAATTTGGTGTGCTATTCATGATGCTCAGGAG 720
 QY 721 CCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGTCCAT 780
 Db 721 CCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGTCCAT 780
 QY 781 CCGGACCATCTACTGCTGTGTGGCTCTTCAACCTCTGTTTGTGCTTCCATATCAC 840
 Db 781 CCGGACCATCTACTGCTGTGTGGCTCTTCAACCTCTGTTTGTGCTTCCATATCAC 840
 QY 841 TCGCTCTCTTACCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGTCTTTGATGGC 900
 Db 841 TCGCTCTCTTACCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGTCTTTGATGGC 900
 QY 901 AG-CCAGTGTGGCTACAGATATGAGAGGCTCTGTGAGTGTGAGAGTGTGCTCAACC 959
 Db 901 AGCCAGTGTGGCTACAGATATGAGAGGCTCTGTGAGTGTGAGAGTGTGCTCAACC 960
 QY 960 CAGTCTGTACTTTCTTTCAAGGGGGCAAAATAGAGTCAAGTCTCTCCAGAACTGAG 1018
 Db 961 CAGTCTGTACTTTCTTTCAAGGGGGCAAAATAGAGTCAAGTCTCTCCAGAACTGAG 1019

```
BD187417
LOCUS BD187417 1020 bp DNA linear PAT 17-JUL-2003
DEFINITION Novel polypeptide.
ACCESSION BD187417
VERSION BD187417.1 GI:32997156
KEYWORDS JP 2003009885-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1020)
AUTHORS Fidoock, M.D.
TITLE Novel polypeptide
JOURNAL Patent: JP 2003009885-A 2 14-JAN-2003;
Pfizer Ltd (EP/GB only), Pfizer Inc (US JP EP except GB)
COMMENT OS Homo sapiens
PN JP 2003009885-A/2
PD 14-JAN-2003
PF 17-DEC-2001 JP 2001382707
PR 18-DEC-2000 GB 0030855.1, 17-JAN-2001 GB 0101222.8 PI
mark david fidoock
CC
FH Key Location/Qualifiers.
FEATURES
source
1..1020
Location/Qualifiers.
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 92.7%; Score 1002.8; DB 6; Length 1020;
Best Local Similarity 99.6%; Pred. No. 1.5e-246;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 1 ATGCTGCCATTTTGTCTTCCAGGGGAAGCAGAGGGGAGCGTGGTGGAGCTCTG 60
DB 1 ATGCTGCCATTTTGTCTTCCAGGGGAAGCAGAGGGGAGCGTGGTGGAGCTCTG 60
QY 61 CTCTGGAGGAGCGTCCCGGAGCATGGAGAAGGTGGACATGAATACATCACAGGAACAA 120
DB 61 CTCTGGAGGAGCGTCCCGGAGCATGGAGAAGGTGGACATGAATACATCACAGGAACAA 120
QY 121 GGTCTGCGAGTCTCAGAGAAGTGAAGAGGTGGACATGAATACATCACAGGAACAA 180
DB 121 GGTCTGCGAGTCTCAGAGAAGTGAAGAGGTGGACATGAATACATCACAGGAACAA 180
QY 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTTGTGGCACTCTCTGGGGCCAAACC 240
DB 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTTGTGGCACTCTCTGGGGCCAAACC 240
```

```
QY 601 TATGACATGACGACGCAAGAGAAATTTGATCGGCTTTTTCGCTACGGCATAGTTCTGACA 660
DB 601 TATGACATGACGACGCAAGAGAAATTTGATCGGCTTTTTCGCTACGGCATAGTTCTGACA 660
QY 661 TTGCTGCGCTTTTCTTT--CCCTCTTGGTCAATTTTGGTGTCTAFTCACTGATGCTCAGG 718
DB 661 TTGCTGCGCTTTTCTTTCCCTCTTGGTCAATTTTGGTGTCTAFTCACTGATGCTCAGG 720
QY 719 AGCTGATCAGCCAGAGGAGAACTCATGAGGACAGCAACACAGCCGAGCCAGGTCC 778
DB 721 AGCTGATCAGCCAGAGGAGAACTCATGAGGACAGCAACACAGCCGAGCCAGGTCC 780
QY 779 ATCCGACCATCTACTGCTGTGGCTCTTCACTCCCTCTGTTTGTGCGCTTCCATATC 838
DB 781 ATCCGACCATCTACTGCTGTGGCTCTTCACTCCCTCTGTTTGTGCGCTTCCATATC 840
QY 839 ACTGCTCTCTTCTACTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATG 898
DB 841 ACTGCTCTCTTCTACTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATG 900
QY 899 GCAGCCAGTGTGGCTACAAGATATGGAGGCTCTCTGCTGAGTGTGAGCAGCTGCTCAAC 958
DB 901 GCAGCCAGTGTGGCTACAAGATATGGAGGCTCTCTGCTGAGTGTGAGCAGCTGCTCAAC 960
QY 959 CGAGTCTGTACTTTTCAAGGGGGGCAAAAATAGAGTCAGGCTCTCTCCAGAAACTGA 1018
DB 961 CGAGTCTGTACTTTTCAAGGGGGGCAAAAATAGAGTCAGGCTCTCTCCAGAAACTGA 1020
```

```
RESULT 13
AX458308
LOCUS AX458308 1020 bp DNA linear PAT 08-JUL-2002
DEFINITION Sequence 3 from Patent EP1215214.
ACCESSION AX458308
VERSION AX458308.1 GI:21725002
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Fidoock, M.D.
TITLE Novel polypeptide
JOURNAL Patent: EP 1215214-A 3 19-JUN-2002;
Pfizer Limited (GB); PFIZER INC. (US)
FEATURES
source
1..1020
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
```

```
Query Match 92.7%; Score 1002.8; DB 6; Length 1020;
Best Local Similarity 99.6%; Pred. No. 1.5e-246;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
QY 1 ATGCTGCCATTTTGTCTTCCAGGGGAAGCAGAGGGGAGCGTGGTGGAGCTCTG 60
DB 1 ATGCTGCCATTTTGTCTTCCAGGGGAAGCAGAGGGGAGCGTGGTGGAGCTCTG 60
QY 61 CTCTGGAGGAGCGTCCCGGAGCATGGAGAAGGTGGACATGAATACATCACAGGAACAA 120
DB 61 CTCTGGAGGAGCGTCCCGGAGCATGGAGAAGGTGGACATGAATACATCACAGGAACAA 120
QY 121 GGTCTGCGAGTCTCAGAGAAGTGAAGAGGTGGACATGAATACATCACAGGAACAA 180
DB 121 GGTCTGCGAGTCTCAGAGAAGTGAAGAGGTGGACATGAATACATCACAGGAACAA 180
QY 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTTGTGGCACTCTCTGGGGCCAAACC 240
DB 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTTGTGGCACTCTCTGGGGCCAAACC 240
```

241 AAGCGTGGAGCTGTGCCAACCTATCTGTGTGAACCTGATGTGGCCGACCTGCTTTAT 300
 Db AAGCGTGGAGCTGTGCCAACCTATCTGTGTGAACCTGATGTGGCCGACCTGCTTTAT 300
 QY GTGCTATTGCCCTTCCCTATCATCATCACTACTCTAGATGACAGGTGGCCCTTCGGGGAG 360
 Db GTGCTATTGCCCTTCCCTATCATCATCACTACTCTAGATGACAGGTGGCCCTTCGGGGAG 360
 QY CTGCTCTGCAAGCTGTGCACTTCTCTGTATATCAACCTTTACGGCAGCATCTCTGCTG 420
 Db CTGCTCTGCAAGCTGTGCACTTCTCTGTATATCAACCTTTACGGCAGCATCTCTGCTG 420
 QY CTGACCTGCATCTCTGTGCAACCTTCTAGTGTGTGCGACCACTGTGCTGCTGCC 480
 Db CTGACCTGCATCTCTGTGCAACCTTCTAGTGTGTGCGACCACTGTGCTGCTGCC 480
 QY TACCGGACCCGACGATGCTGTGCGACCACTGCGGACCACTGCGGACCTGCTGCTC 540
 Db TACCGGACCCGACGATGCTGTGCGACCACTGCGGACCACTGCGGACCTGCTGCTC 540
 QY CAGTGTGCTGCACTGTGCGACCTTCTGCGACCACTGCGGACCACTGCGGACCTGCTG 600
 Db CAGTGTGCTGCACTGTGCGACCTTCTGCGACCACTGCGGACCACTGCGGACCTGCTG 600
 QY TATGACATGACCAAGCAAGATTTTGTGCGCTTTTGTGCGCTTTTGTGCGCTTTTGTGCA 660
 Db TATGACATGACCAAGCAAGATTTTGTGCGCTTTTGTGCGCTTTTGTGCGCTTTTGTGCA 660
 QY TTGCTGCTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 718
 Db TTGCTGCTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
 QY AGCTGATCAAGCAGAGGAGCACTCATGAGGACCAAGCAAGCAAGCAAGCAAGCAAGCTC 778
 Db AGCTGATCAAGCAGAGGAGCACTCATGAGGACCAAGCAAGCAAGCAAGCAAGCAAGCTC 780
 QY ATCCGACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 838
 Db ATCCGACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 840
 QY ACTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 898
 Db ACTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
 QY GCAGCAGTGTGCGCTTCAAGATATGAGGCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 958
 Db GCAGCAGTGTGCGCTTCAAGATATGAGGCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 960
 QY CCAGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1018
 Db CCAGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1020

RESULT 14
 AX375230 1076 bp DNA linear PAT 01-MAR-2002
 LOCUS Sequence 20 from Patent WO0210387.
 DEFINITION AX375230
 ACCESSION AX375230.1 GI:19169948
 VERSION
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R.,
 Elliot, V.S., Ramkumar, J., Baughn, M.R., Kallik, D.A., Walla, N.K.,
 Hafalia, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L.,
 Kearney, B., Graul, R.C., Warren, B.A. and Ding, L.
 G-protein coupled receptors
 Patent: WO 0210387-A 20 07-FEB-2002;
 Inocyte Genomics, Inc. (US)
 TITLE Location/Qualifiers
 JOURNAL
 FEATURES

source 1..1076
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Inocyte ID No: 7474806CBI"

ORIGIN

Query Match 92.7%; Score 1002.8; DB 6; Length 1076;
 Best Local Similarity 99.6%; Pred. No. 1.5e-246;
 Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1 ATGCTGTCCATTTTCTTCTTCCAGGGAAGCAGAGCGGAGCGCTGCGGAGCTCTG 60
 Db 57 ATGCTGTCCATTTTCTTCTTCCAGGGAAGCAGAGCGGAGCGCTGCGGAGCTCTG 116
 QY 61 CTCTTGGAGGAGCTTCCCGGACATGGAGAGGTGGACATGAATATACACAGAAACAA 120
 Db 117 CTCTTGGAGGAGCTTCCCGGACATGGAGAGGTGGACATGAATATACACAGAAACAA 176
 QY 121 GGTCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 180
 Db 177 GGTCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 236
 QY 181 ATCTTTATCTTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240
 Db 237 ATCTTTATCTTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 296
 QY 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCGGACCTGCTTTAT 300
 Db 297 AAGCGTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCGGACCTGCTTTAT 356
 QY 301 GTGCTATTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360
 Db 357 GTGCTATTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 416
 QY 361 CTGCTCTGCAAGCTGTGCACTTCTGTTCTATATCAACCTTTACGGGAGCATCTGCTG 420
 Db 417 CTGCTCTGCAAGCTGTGCACTTCTGTTCTATATCAACCTTTACGGGAGCATCTGCTG 476
 QY 421 CTGACCTGCATCTCTGTGCAACCTTCTAGTGTGTGCAACCTTCTGCTGCTGCTG 480
 Db 477 CTGACCTGCATCTCTGTGCAACCTTCTAGTGTGTGCAACCTTCTGCTGCTGCTG 536
 QY 481 TACCGGACCCGACGATGCTGTGCGACCACTGCGGACCACTGCGGACCTGCTGCTC 540
 Db 537 TACCGGACCCGACGATGCTGTGCGACCACTGCGGACCACTGCGGACCTGCTGCTC 596
 QY 541 CAGTGTGCTGCACTTCCCGGACATGGAGAGGTGGACATGAATATACACAGAAACAA 600
 Db 597 CAGTGTGCTGCACTTCCCGGACATGGAGAGGTGGACATGAATATACACAGAAACAA 656
 QY 601 TATGACATGACCAAGCAAGATTTTGTGCGCTTTTGTGCGCTTTTGTGCGCTTTTGTGCA 660
 Db 657 TATGACATGACCAAGCAAGATTTTGTGCGCTTTTGTGCGCTTTTGTGCGCTTTTGTGCA 716
 QY 661 TTGCTGCTGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 718
 Db 717 TTGCTGCTGCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 776
 QY 719 AGCTGATCAAGCAGAGGAGCACTCATGAGGACCAAGCAAGCAAGCAAGCAAGCAAGCTC 778
 Db 777 AGCTGATCAAGCAGAGGAGCACTCATGAGGACCAAGCAAGCAAGCAAGCAAGCAAGCTC 836
 QY 779 ATCCGACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 838
 Db 837 ATCCGACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 896
 QY 839 ACTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 898
 Db 897 ACTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 956
 QY 899 GCAGCAGTGTGCGCTTCAAGATATGAGGCTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAG 958

```

Db      957 GCACCCAGTGTGGCTACAGATATGAGGCGCTCTGGTGAAGTGAGCAGCTGCCTCAAC 1016
QY      959 CAGTCTCTGACTTTCTTCAAGGGGGCAAAAATAGATCAGGCTCTCCAGAACTGA 1018
Db      1017 CCAGTCTCTGACTTTCTTCAAGGGGGCAAAAATAGATCAGGCTCTCCAGAACTGA 1076

RESULT 15
BD095704
LOCUS   1002 bp      DNA      linear      PAT 27-AUG-2002
DEFINITION
Novel guanosine triphosphate-bound protein-coupled receptors and
genes encoding them, and their production and use.
ACCESSION
BD095704.1 GI:22641292
VERSION
WO 0148188-A/16.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1002)
AUTHORS
Matsumoto,S., Oda,T., Saito,Y., Noriyuki, Morikawa, Yoshida,K.,
Suwa,M., Sugiyama,T., Kishimoto,T., Kanzaki,K., Yasuda,S. and
Inoue,Y.
TITLE
Novel guanosine triphosphate-bound protein-coupled receptors and
genes encoding them, and their production and use
JOURNAL
Patent: WO 0148188-A 16 05-JUL-2001;
HELIX RESEARCH INSTITUTE,SHUNICHIRO MATSUMOTO,TAMAKI ODA,YOKO
SAITO, NORIYUKI MORIKAWA,KENJI YOSHIDA, MAKIKO SUWA,TOMOYASU
SUGIYAMA, TOSHIMITSU KISHIMOTO,KOJI KANZAKI,SHINICHIRO
YASUDA,YOSHIHISA INOUE
COMMENT
OS Homo sapiens (human)
PN WO 0148188-A/16
PD 05-JUL-2001
PF 28-DEC-2000 WO 2000JP009408
PR 28-DEC-1999 JP 99P 375152,31-MAR-2000 JP OOP 101339 PI
SHUNICHIRO MATSUMOTO,TAMAKI ODA,YOKO SAITO,NORIYUKI PI
MORIKAWA,KENJI YOSHIDA,
PI MAKIKO SUWA,TOMOYASU SUGIYAMA,TOSHIMITSU KISHIMOTO,KOJI PI
KANZAKI,
PI SHINICHIRO YASUDA,YOSHIHISA INOUE
PC C12N15/09,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C07K14/705, PC
C07K16/28,
PC C12P21/02,C12Q1/02,C12Q1/68,A61K31/711,A61K48/00,A61P43/00, PC
G01N33/15,
PC G01N33/50
CC Novel guanosine triphosphate-bound protein-coupled receptors
and genes
CC encoding them, and their production and use
FH Key Location/Qualifiers
FT source 1..1002
FT /organism="Homo sapiens (human)".
FEATURES
Source
1..1002
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 92.1%; Score 996.4; DB 6; Length 1002;
Best Local Similarity 99.9%; Pred. No. 6.6e-245;
Matches 997; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      85 ATGGAGAGGTGGACATACATACACAGGACAGGCTCTGCCAGTTCTCAGAGAG 144
Db      1 ATGGAGAGGTGGACATACATACACAGGACAGGCTCTGCCAGTTCTCAGAGAG 60

QY      145 TACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTTTATCTAGGGCTGCCACTA 204
Db      61 TACAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTTTATCTAGGGCTGCCACTA 120

QY      205 AATGGCACTGTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACACC 264
Db      121 AATGGCACTGTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACACC 180

```

Search completed: November 10, 2004, 10:44:07
Job time : 4777 secs

```

QY      265 TATCTGTGTGACCTGATGTTGGCGGACCTGCTTTATGTGCTATTGGCCCTTCCTCATCATC 324
Db      181 TATCTGTGTGACCTGATGTTGGCGGACCTGCTTTATGTGCTATTGGCCCTTCCTCATCATC 240
QY      325 ACCTACTCACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTC 384
Db      241 ACCTACTCACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTC 300
QY      385 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCATCTCTGTGCACCAG 444
Db      301 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCATCTCTGTGCACCAG 360
QY      445 TTCCTAGTGTGTGCCACCCACTCTGTTGCTGCTACCGGACCCGCGCAGCATGCTCTGG 504
Db      361 TTCCTAGTGTGTGCCACCCACTCTGTTGCTGCTACCGGACCCGCGCAGCATGCTCTGG 420
QY      505 CTGGGCACAGCACACCTGGGCCCTGGTGGTCTCCAGTCTGCTGCCACACTGGCCCTTC 564
Db      421 CTGGGCACAGCACACCTGGGCCCTGGTGGTCTCCAGTCTGCTGCCACACTGGCCCTTC 480
QY      565 TCCCACACGGGACTACATCAATGGCCAGATGATCTGGTATGACATGACACGACCAAGAGAAT 624
Db      481 TCCCACACGGGACTACATCAATGGCCAGATGATCTGGTATGACATGACACGACCAAGAGAAT 540
QY      625 TTTGATCGGCTTTTGGCTACGGCATAGTTCTGACATGTTCTGGCTTTCTTTTCCCTCTT 684
Db      541 TTTGATCGGCTTTTGGCTACGGCATAGTTCTGACATGTTCTGGCTTTCTTTTCCCTCTT 600
QY      685 GGTCAATTTGGTGTGCTATTCACTGATGTTCTGAGGAGCTGATCAAGCCAGAGGAGAACCT 744
Db      601 GGTCAATTTGGTGTGCTATTCACTGATGTTCTGAGGAGCTGATCAAGCCAGAGGAGAACCT 660
QY      745 CATGAGGACAGGCAACACAGCCCGAGCCAGGTCCATCCGACCACTCTACTTGTGTGTGG 804
Db      661 CATGAGGACAGGCAACACAGCCCGAGCCAGGTCCATCCGACCACTCTACTTGTGTGTGG 720
QY      805 CCTCTTCACTCTGTTTGTGGCTTCCATATCACTGCTGCTCTTACCTCACCACCTG 864
Db      721 CCTCTTCACTCTGTTTGTGGCTTCCATATCACTGCTGCTCTTACCTCACCACCTG 780
QY      865 CTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCTTCAAGATATG 924
Db      781 CTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCTTCAAGATATG 840
QY      925 GAGGCTCTGTGTGAGTGTGAGCAGCTGCTCAACCCAGTCTCTGACTTTCTTTCAAGGGG 984
Db      841 GAGGCTCTGTGTGAGTGTGAGCAGCTGCTCAACCCAGTCTCTGACTTTCTTTCAAGGGG 900
QY      985 GGCAGAAATAGAGTCAGGCTCTCCAGAACTGAGGCGAGAACAAAGTTGGGTGAGCATCCA 1044
Db      901 GGCAGAAATAGAGTCAGGCTCTCCAGAACTGAGGCGAGAACAAAGTTGGGTGAGCATCCA 960
QY      1045 GCTGGGAGGAGAGATGCCAGGGTTGAACAGATCTGG 1082
Db      961 GCTGGGAGGAGAGATGCCAGGGTTGAACAGATCTGG 998

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 09:14:47 ; Search time 611 Seconds
(without alignments)
9534.556 Million cell updates/sec

Title: US-10-763-972-1
Perfect score: 1082
Sequence: 1 agcgttcatttgcctcc.....ccagggtgaacagatctgg 1082

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
- 20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
- 21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1082	100.0	1082	15	US-10-023-586B-1
2	1082	100.0	1082	17	US-10-763-972-1
3	1082	100.0	1334	15	US-10-017-161-707
4	1082	100.0	1334	15	US-10-292-798-619
5	1082	100.0	3143	15	US-10-275-910-1
6	1059	97.9	1108	16	US-10-072-012-165
7	1006	93.0	1017	10	US-09-885-453-3
8	1006	93.0	1017	15	US-10-079-384-5
9	1002.8	92.7	1020	15	US-10-023-586B-3
10	1002.8	92.7	1020	17	US-10-763-972-3
11	1002.8	92.7	1076	16	US-10-333-946-20
12	996.4	92.1	1002	15	US-10-088-726-25

13	851	78.7	851	15	US-10-275-910-6	Sequence 6, Appli
14	850	78.6	850	15	US-10-275-910-4	Sequence 4, Appli
15	510	47.1	510	10	US-09-782-974C-13	Sequence 13, Appli
c 16	396.4	36.6	681	11	US-09-801-944B-47	Sequence 47, Appli
17	370	34.2	585	11	US-09-801-944B-56	Sequence 56, Appli
18	227.8	21.1	2025	15	US-10-225-567A-216	Sequence 216, App
19	227.8	21.1	2025	15	US-10-101-510-722	Sequence 722, App
20	227.8	21.1	2025	16	US-10-305-720-1482	Sequence 1482, Ap
21	227.8	21.1	2025	17	US-10-776-827-74	Sequence 74, Appli
22	226.2	20.9	2118	15	US-10-101-510-431	Sequence 431, App
c 23	222	20.5	551	10	US-09-791-932-32	Sequence 32, Appli
24	178.6	16.5	1098	15	US-10-225-567A-331	Sequence 331, App
25	178.6	16.5	1429	10	US-09-077-173A-1	Sequence 1, Appli
26	178.6	16.5	1429	16	US-10-305-720-1068	Sequence 1068, Ap
27	178.6	16.5	1429	17	US-10-753-695-1	Sequence 1, Appli
28	173.8	16.1	1651	15	US-10-366-288-41	Sequence 41, Appli
29	170	15.7	234	11	US-09-801-944B-55	Sequence 55, Appli
30	164.8	15.2	984	15	US-10-278-087A-57	Sequence 57, Appli
31	163.2	15.1	1571	15	US-10-354-358-9	Sequence 9, Appli
32	163.2	15.1	1571	16	US-10-305-720-1108	Sequence 1108, Ap
33	163.2	15.1	1832	15	US-10-225-567A-222	Sequence 222, App
34	163.2	15.1	1832	15	US-10-172-116-994	Sequence 994, App
35	163.2	15.1	1832	15	US-10-295-027-1071	Sequence 1071, Ap
36	163.2	15.1	1832	16	US-10-342-887-994	Sequence 994, App
37	158.4	14.6	1922	16	US-10-152-319A-2068	Sequence 2068, Ap
38	156.8	14.5	984	15	US-10-242-499-1	Sequence 1, Appli
39	148.8	13.8	984	15	US-10-278-087A-41	Sequence 41, Appli
40	111.8	10.3	1104	15	US-10-041-615-33	Sequence 33, Appli
41	105.8	9.8	2424	15	US-10-225-567A-218	Sequence 218, Appli
42	97.8	9.0	1543	10	US-09-891-138A-1	Sequence 1, Appli
43	97.4	9.0	1119	10	US-09-826-509-538	Sequence 538, App
44	97.4	9.0	1773	14	US-10-112-599A-3	Sequence 3, Appli
45	97.4	9.0	1773	15	US-10-225-567A-107	Sequence 107, App

ALIGNMENTS

RESULT 1

US-10-023-586B-1
; Sequence 1, Application US/10023586B
; Publication No. US20030166882A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (EP (GB) only)
; APPLICANT: Pfizer Inc. (US, JP, EP except GB)
; APPLICANT: Fidock, Mark David
; TITLE OF INVENTION: No. US20030166882A1e1 Polypeptide
; FILE REFERENCE: PC10960AGPR
; CURRENT APPLICATION NUMBER: US/10/023,586B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030855.1
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,563
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/265,688
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: GB 0101222.8
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-586B-1

Query Match 100.0%; Score 1082; DB 15; Length 1082;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCATTTCCTTCCTTCAGGGGAGCAGAGCGGCGCTGAGCTCTG 60
|||||
Db 1 ATGCTGTCATTTCCTTCCTTCAGGGGAGCAGAGCGGCGCTGAGCTCTG 60
|||||

```
QY 61 CTCCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATGATATACATCACAGGACAA 120
Db 61 CTCCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATGATATACATCACAGGACAA 120
QY 121 GGTCTCTGCCAGTCTCTCAGAGAAGTACAAAGAGTCTACCTCTCCCTGGGCTACAGTATC 180
Db 121 GGTCTCTGCCAGTCTCTCAGAGAAGTACAAAGAGTCTACCTCTCCCTGGGCTACAGTATC 180
QY 181 ATCTTTATCCTAGGCTGCCACTAAATGGGCACTGTCTTGTGGCACTCTCTGGGGCAAACC 240
Db 181 ATCTTTATCCTAGGCTGCCACTAAATGGGCACTGTCTTGTGGCACTCTCTGGGGCAAACC 240
QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTAACCTGTAGTGGCGGACCTGCTTTAT 300
Db 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTAACCTGTAGTGGCGGACCTGCTTTAT 300
QY 301 GTGCTATGGCCCTTCTCATCATCACCTACTCATAGATGACAGGTGGCCCTTGGGGAG 360
Db 301 GTGCTATGGCCCTTCTCATCATCACCTACTCATAGATGACAGGTGGCCCTTGGGGAG 360
QY 361 CTGCTCTGCAAGCTGGTGCATCTCTGTCTTATATCAACCTTTACGCGACATCTGCTG 420
Db 361 CTGCTCTGCAAGCTGGTGCATCTCTGTCTTATATCAACCTTTACGCGACATCTGCTG 420
QY 421 CTGACCTGCACTCTCTGTGCAACAGTCTCTAGGTGTGGCCACCACTGTCTGCTGCC 480
Db 421 CTGACCTGCACTCTCTGTGCAACAGTCTCTAGGTGTGGCCACCACTGTCTGCTGCC 480
QY 481 TACCGAGCCCGAGGATGCTGCTGGGACACAGCACCACTGGGCGCTGTGTGCTTC 540
Db 481 TACCGAGCCCGAGGATGCTGCTGGGACACAGCACCACTGGGCGCTGTGTGCTTC 540
QY 541 CAGCTGTGCCACACCTGCGCTTCTCCACACGAGCTACATCAATGCGCCAGATGATCTGG 600
Db 541 CAGCTGTGCCACACCTGCGCTTCTCCACACGAGCTACATCAATGCGCCAGATGATCTGG 600
QY 601 TATGACATGACACGACCAAGAGATTTTGTATCGGCTTTTTCCTACGCACTAGTCTGACA 660
Db 601 TATGACATGACACGACCAAGAGATTTTGTATCGGCTTTTTCCTACGCACTAGTCTGACA 660
QY 661 TTGCTGTGCTTTCTTCTCCCTCTGGTCAATTTTGTGTGCTATTTCACTGATGTCAGGAG 720
Db 661 TTGCTGTGCTTTCTTCTCCCTCTGGTCAATTTTGTGTGCTATTTCACTGATGTCAGGAG 720
QY 721 CCTGATCAAGCCAGAGGAACTCTATGAGGACAGGCAACACAGCCCGAGCCAGGTCCAT 780
Db 721 CCTGATCAAGCCAGAGGAACTCTATGAGGACAGGCAACACAGCCCGAGCCAGGTCCAT 780
QY 781 CCGGACCATCTACTGTGTGTGGCTCTTCCCTCTGTTTGTGCTTTGTGCTTCCATATCAC 840
Db 781 CCGGACCATCTACTGTGTGTGGCTCTTCCCTCTGTTTGTGCTTTGTGCTTCCATATCAC 840
QY 841 TCGCTCTCTTCTACCTACCATCTGCTTTCTGCTTTCTCAGGACTGCGCAGCTTTGATGGC 900
Db 841 TCGCTCTCTTCTACCTACCATCTGCTTTCTGCTTTCTCAGGACTGCGCAGCTTTGATGGC 900
QY 901 AGCCAGTGTGGCTTACAGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAACCC 960
Db 901 AGCCAGTGTGGCTTACAGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAACCC 960
QY 961 AGTCTCTGCTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCCAGAACTGAGG 1020
Db 961 AGTCTCTGCTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCCAGAACTGAGG 1020
QY 1021 CAGAACAGTGTGGTGAAGATTCAGCTGGGAGGAGAGATGCCAGGGTTGAACATCT 1080
Db 1021 CAGAACAGTGTGGTGAAGATTCAGCTGGGAGGAGAGATGCCAGGGTTGAACATCT 1080
QY 1081 GG 1082
Db 1081 GG 1082
```

```
RESULT 2
US-10-763-972-1
; Sequence 1, Application US/10763972
; Publication NO. US20040137500A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER INC.
; APPLICANT: Fidoek, Mark David
; TITLE OF INVENTION: Novel Polypeptide
; FILE REFERENCE: PC10960B
; CURRENT APPLICATION NUMBER: US/10/763,972
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: GB 0030855.1
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,563
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/265,688
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: GB 0101222.8
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-763-972-1

Query Match 100.0%; Score 1082; DB 17; Length 1082;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTGCATTTTGTCTTCCAGGGGAAGCAGAGCGGAGCCGTCTGGAGCTGTG 60
Db 1 ATGCTGTGCATTTTGTCTTCCAGGGGAAGCAGAGCGGAGCCGTCTGGAGCTGTG 60
QY 61 CTCCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATGATATACATCACAGGACAA 120
Db 61 CTCCTGGAGGAGCCTCCGGGACATGGAGAGGTGGACATGATATACATCACAGGACAA 120
QY 121 GGTCTCTGCCAGTCTCTCAGAGAAGTACAAAGAGTCTACCTCTCCCTGGGCTACAGTATC 180
Db 121 GGTCTCTGCCAGTCTCTCAGAGAAGTACAAAGAGTCTACCTCTCCCTGGGCTACAGTATC 180
QY 181 ATCTTTATCCTAGGCTGCCACTAAATGGGCACTGTCTTGTGGCACTCTCTGGGGCAAACC 240
Db 181 ATCTTTATCCTAGGCTGCCACTAAATGGGCACTGTCTTGTGGCACTCTCTGGGGCAAACC 240
QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTAACCTGTAGTGGCGGACCTGCTTTAT 300
Db 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTAACCTGTAGTGGCGGACCTGCTTTAT 300
QY 301 GTGCTATGGCCCTTCTCATCATCACCTACTCATAGATGACAGGTGGCCCTTGGGGAG 360
Db 301 GTGCTATGGCCCTTCTCATCATCACCTACTCATAGATGACAGGTGGCCCTTGGGGAG 360
QY 361 CTGCTCTGCAAGCTGGTGCATCTCTGTCTTATATCAACCTTTACGCGACATCTGCTG 420
Db 361 CTGCTCTGCAAGCTGGTGCATCTCTGTCTTATATCAACCTTTACGCGACATCTGCTG 420
QY 421 CTGACCTGCACTCTCTGTGCAACAGTCTCTAGGTGTGGCCACCACTGTCTGCTGCC 480
Db 421 CTGACCTGCACTCTCTGTGCAACAGTCTCTAGGTGTGGCCACCACTGTCTGCTGCC 480
QY 481 TACCGAGCCCGAGGATGCTGCTGGGACACAGCACCACTGGGCGCTGTGTGCTTC 540
Db 481 TACCGAGCCCGAGGATGCTGCTGGGACACAGCACCACTGGGCGCTGTGTGCTTC 540
QY 541 CAGCTGTGCCACACCTGCGCTTCTCCACACGAGCTACATCAATGCGCCAGATGATCTGG 600
Db 541 CAGCTGTGCCACACCTGCGCTTCTCCACACGAGCTACATCAATGCGCCAGATGATCTGG 600
QY 601 TATGACATGACACGACCAAGAGATTTTGTATCGGCTTTTTCCTACGCACTAGTCTGACA 660
```

Db	601	TATGATGACGACGACGAGAAATTTTATCGGCTTTTTCGCTACGGCATAGTTCTTGACA	660	117	ATGCTGTCCATTTTGTCTTCTTCCAGGGGAGCAGAGAGCGGAGCCCTGCTGGAGCTCTG	176
Qy	661	TTGTCCTGGCTTTCTTCCCTCCTTGTGTCATTTTGGTGTGCTATTACTATGTCAGGAG	720	61	CTCCTGGAGGAGCCTCCGGGACATGAGAGAGTGGACATGAATACATACACAGAACAA	120
Db	661	TTGTCCTGGCTTTCTTCCCTCCTTGTGTCATTTTGGTGTGCTATTACTATGTCAGGAG	720	177	CTCCTGGAGGAGCCTCCGGGACATGAGAGAGTGGACATGAATACATACACAGAACAA	236
Qy	721	CCTGATCAAGCCAGAGAGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGTCCAT	780	121	GCTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCTCCCTGGCCCTACAGTATC	180
Db	721	CCTGATCAAGCCAGAGAGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGTCCAT	780	237	GCTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCTCCCTGGCCCTACAGTATC	296
Qy	781	COGGAACATCCTACTGCTGGTGGCCCTCTTACCCCTCTGTTTGTGGCCCTTCCATATCAC	840	181	ATCTTTATCTAGGCTGCCACTAAATGGGACTCTCTTGTGGCACTCTCTGGGGCCAAACC	240
Db	781	COGGAACATCCTACTGCTGGTGGCCCTCTTACCCCTCTGTTTGTGGCCCTTCCATATCAC	840	297	ATCTTTATCTAGGCTGCCACTAAATGGGACTCTCTTGTGGCACTCTCTGGGGCCAAACC	356
Qy	841	TGCTCTCTTCTACTCCTACCATCTGCTTTCTGCTTTCTCAGGAGTGCAGCTCTTTGATGC	900	241	AAGGCTGAGCTGTGCCACCACCTATCTGCTGAACCTGATGCTGGCGGACCTCTCTTTAT	300
Db	841	TGCTCTCTTCTACTCCTACCATCTGCTTTCTGCTTTCTCAGGAGTGCAGCTCTTTGATGC	900	357	AAGGCTGAGCTGTGCCACCACCTATCTGCTGAACCTGATGCTGGCGGACCTCTCTTTAT	416
Qy	901	AGCCAGTGTGGCTACAGATATGAGGCTCTGCTGAGTGTGAGCGTGCCTCAACC	960	301	GTGCTATTTGCCCTTCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	360
Db	901	AGCCAGTGTGGCTACAGATATGAGGCTCTGCTGAGTGTGAGCGTGCCTCAACC	960	417	GTGCTATTTGCCCTTCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	476
Qy	961	AGTCCTGTACTTTTTCAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTGAGG	1020	361	CTGCTCTGCAAGCTGGTGCACCTTCTGCTTCTATATCAACCTTTTACGGCAGCATCTGCTG	420
Db	961	AGTCCTGTACTTTTTCAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTGAGG	1020	477	CTGCTCTGCAAGCTGGTGCACCTTCTGCTTCTATATCAACCTTTTACGGCAGCATCTGCTG	536
Qy	1021	CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT	1080	421	CTGACCTGCATCTCTGTGTCACCATGTTCTAGGTGTGTGCCACCATGTTGTTCGTCGCC	480
Db	1021	CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT	1080	537	CTGACCTGCATCTCTGTGTCACCATGTTCTAGGTGTGTGCCACCATGTTGTTCGTCGCC	596
Qy	1081	GG 1082		481	TACGGACCCGAGGATGCTGGTGGGACACGACCACTGGGCGCCCTGGTGGTCCCTC	540
Db	1081	GG 1082		597	TACGGACCCGAGGATGCTGGTGGGACACGACCACTGGGCGCCCTGGTGGTCCCTC	656

RESULT 3
US-10-017-161-707
; Sequence 707, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ASURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 707
; LENGTH: 1334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1334)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(715)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (834)..(1134)
US-10-017-161-707
Query Match 100.0%; Score 1082; DB 15; Length 1334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCTGTCCATTTTGTCTTCTTCCAGGGGAGCAGAGAGCGGAGCCCTGCTGGAGCTCTG 60

Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGCTGTCCATTTTGTCTTCCTCCAGGGGAAGCAGAGGGGAGCGTGTGGAGCTCTG	60
Db	436	ATGCTGTCCATTTTGTCTTCCTCCAGGGGAAGCAGAGGGGAGCGTGTGGAGCTCTG	495
Qy	61	CTCTGGAGGAGCCCTCCGGGACATGGAGAGGTGGACATGAATACATCAGGAACAA	120
Db	496	CTCTGGAGGAGCCCTCCGGGACATGGAGAGGTGGACATGAATACATCAGGAACAA	555
Qy	121	GGTCTCTGCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC	180
Db	556	GGTCTCTGCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC	615
Qy	181	ATCTTTATCTAGGGTGGCCATTAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC	240
Db	616	ATCTTTATCTAGGGTGGCCATTAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC	675
Qy	241	AAGGGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGTGGCCGACCTGTTAT	300
Db	676	AAGGGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGTGGCCGACCTGTTAT	735
Qy	301	GTGCTATTGGCCCTTCCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG	360
Db	736	GTGCTATTGGCCCTTCCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG	795
Qy	361	CTGCTCTGGAAGCTGTGGTGCATCTCTCTGTCTTATATCAACTTTACGGCAGCATCTGTCTG	420
Db	796	CTGCTCTGGAAGCTGTGGTGCATCTCTCTGTCTTATATCAACTTTACGGCAGCATCTGTCTG	855
Qy	421	CTGACCTCGATCTCTGTGACACCACTCTAGGTGTGTGCCACCACTGTGTCTGGTGCCTC	480
Db	856	CTGACCTCGATCTCTGTGACACCACTCTAGGTGTGTGCCACCACTGTGTCTGGTGCCTC	915
Qy	481	TACCGGACCGGAGGATGCTGGTGGGACACGACCACTGGGCCCTGGTGGTCCCTC	540
Db	916	TACCGGACCGGAGGATGCTGGTGGGACACGACCACTGGGCCCTGGTGGTCCCTC	975
Qy	541	CAGTCTGTGCCACACCTGGCCCTCTCCCAACGAGTACATCAATGCCAGATGATCTCG	600
Db	976	CAGTCTGTGCCACACCTGGCCCTCTCCCAACGAGTACATCAATGCCAGATGATCTCG	1035
Qy	601	TATGACATGACCGACCAAGAAATTTTGTATCGGCTTTTGTCCCTACGCGCATAGTTCTGACA	660
Db	1036	TATGACATGACCGACCAAGAAATTTTGTATCGGCTTTTGTCCCTACGCGCATAGTTCTGACA	1095
Qy	661	TTGTCTGGCTTTCTTCCCTCCTTGGTCACTTTTGGTGTGCTATTTCACTGATGTGAGGAG	720
Db	1096	TTGTCTGGCTTTCTTCCCTCCTTGGTCACTTTTGGTGTGCTATTTCACTGATGTGAGGAG	1155
Qy	721	CCTGATCAAGCCAGAGAGAACCTCTAGGACAGGCAACACAGCCCGAGCCAGGTCCAT	780
Db	1156	CCTGATCAAGCCAGAGAGAACCTCTAGGACAGGCAACACAGCCCGAGCCAGGTCCAT	1215
Qy	781	CCGGACCATCTACTGTGTGTGGCCCTTTTACCCCTCTGTTTGTGCCCTTCCATATCAC	840
Db	1216	CCGGACCATCTACTGTGTGTGGCCCTTTTACCCCTCTGTTTGTGCCCTTCCATATCAC	1275
Qy	841	TCGCTCCTTTTACCTACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC	900
Db	1276	TCGCTCCTTTTACCTACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC	1335
Qy	901	AGCCAGTGTGGCCCTACAAGATATGGAGGCCCTCTGGTGAAGTGTGAGCAGCTGCCCAACC	960
Db	1336	AGCCAGTGTGGCCCTACAAGATATGGAGGCCCTCTGGTGAAGTGTGAGCAGCTGCCCAACC	1395
Qy	961	AGTCTCTACTTTCTTTCAAGGGGGGCAAAAATAGACTAGGCTCTCCAGAACTGAGG	1020
Db	1396	AGTCTCTACTTTCTTTCAAGGGGGGCAAAAATAGACTAGGCTCTCCAGAACTGAGG	1455
Qy	1021	CAGAACTGGTGGTGAAGCATCCAGCTGGGAGGAGATGCCCGGTTTGAACAGATCT	1080
Db	1456	CAGAACTGGTGGTGAAGCATCCAGCTGGGAGGAGATGCCCGGTTTGAACAGATCT	1515

Qy 1081 GG 1082
Db 1516 GG 1517

RESULT 6

US-10-072-012-165
; Sequence 165, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Kastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 165
; LENGTH: 1108
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-012-165

Query Match 97.9%; Score 1059; DB 16; Length 1108;
Best Local Similarity 99.8%; Pred.No. 0;
Matches 1081; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 ATGCTGTCCATTTTGTCTTCCTCCAGGGGAAGCAGAGGGGAGCGTGTGGAGCTCTG 60

```

Db 12 ATGCTGTCCATTGTTCTTCTTCCAGGGGAGCAGAGCGGGAGCCGTCTGTGAGCTCTG 71
Qy 61 CTCCTGTGAGGAGGCTCCCGGGACATGGAGAGTGGACATGAATACATCACAGGAACAA 120
Db 72 CTCCTGTGAGGAGGCTCCCGGGACATGGAGAGTGGACATGAATACATCACAGGAACAA 131
Qy 121 GGTCTCTGCCAGTCTCTCAGAGAACTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
Db 132 GGTCTCTGCCAGTCTCTCAGAGAACTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 191
Qy 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGGCACTCTCTGGGCGCAAAACC 240
Db 192 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGGCACTCTCTGGGCGCAAAACC 251
Qy 241 AAGCGCTGGAGCTGTGCCACCACTCTGTGTGAACCTGTATGGTGGCGGACCTGCTTTAT 300
Db 252 AAGCGCTGGAGCTGTGCCACCACTCTGTGTGAACCTGTATGGTGGCGGACCTGCTTTAT 311
Qy 301 GTGCTATGGCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 360
Db 312 GTGCTATGGCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 371
Qy 361 CTGCTCTGCAAGCTGTGTCACCTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
Db 372 CTGCTCTGCAAGCTGTGTCACCTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 431
Qy 421 CTGACCTGTGCATCTGTGTCACCACTCTAGTGTGTGCCACCACTGTGTTGCTGCCCC 480
Db 432 CTGACCTGTGCATCTGTGTCACCACTCTAGTGTGTGCCACCACTGTGTTGCTGCCCC 491
Qy 481 TACCGGACCGCAGGAGCTGTGCTGGCAGCAGCACCCTGGGCGCTGTGTGCTCTC 540
Db 492 TACCGGACCGCAGGAGCTGTGCTGGCAGCAGCACCCTGGGCGCTGTGTGCTCTC 551
Qy 541 CAGCTGTGCTGCACACTGGCCTTCTCCACACGAGCTACATCAATGAGCCAGATGATCTGG 600
Db 552 CAGCTGTGCTGCACACTGGCCTTCTCCACACGAGCTACATCAATGAGCCAGATGATCTGG 611
Qy 601 TATGACATGACCACTGAGAGATTTGATGCGCTTTTGGCTACGGCATAGTCTGACA 660
Db 612 TATGACATGACCACTGAGAGATTTGATGCGCTTTTGGCTACGGCATAGTCTGACA 671
Qy 661 TTGCTGTGCTTTCTTCCCTCTGTGTCATTTTGGTGTGCTATTTCACTGATGCTCAGGAG 720
Db 672 TTGCTGTGCTTTCTTCCCTCTGTGTCATTTTGGTGTGCTATTTCACTGATGCTCAGGAG 730
Qy 721 CCTGATCAAGCCAGAGGAACTCTATGAGGACAGGCAACACAGCCCGAGCCAGGCTCCAT 780
Db 731 CCTGATCAAGCCAGAGGAACTCTATGAGGACAGGCAACACAGCCCGAGCCAGGCTCCAT 790
Qy 781 CCGGACCATCTACTGTGTGTCGCTCTTCAACCTCTGTTTGTGCGCTTCCATATCAC 840
Db 791 CCGGACCATCTACTGTGTGTCGCTCTTCAACCTCTGTTTGTGCGCTTCCATATCAC 850
Qy 841 TCGCTCTCTTACTCTACCACTGCTTTCTGCTTTCTCAGGATGCGAGCTCTGTGATGCG 900
Db 851 TCGCTCTCTTACTCTACCACTGCTTTCTGCTTTCTCAGGATGCGAGCTCTGTGATGCG 910
Qy 901 AG-CCAGCTGTGGCTACAGATATGAGGCGCTCTGCTGCTGTGAGCTGTGAGCTGCTCAACC 959
Db 911 AGCCAGCTGTGGCTACAGATATGAGGCGCTCTGCTGCTGTGAGCTGTGAGCTGCTCAACC 970
Qy 960 CAGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCCTCCAGAAACTGAG 1019
Db 971 CAGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCCTCCAGAAACTGAG 1030
Qy 1020 GCAGAACAGTGTGGTGAAGCATCAGCTGGGAGAGAGATGCGCAGGTTGAACAGATC 1079
Db 1031 GCAGAACAGTGTGGTGAAGCATCAGCTGGGAGAGAGATGCGCAGGTTGAACAGATC 1090
Qy 1080 TGG 1082
Db 1091 TGG 1093

```

```

RESULT 7
US-09-885-453-3
; Sequence 3, Application US/09885453
; Publication No. US2003008080A1
; GENERAL INFORMATION:
; APPLICANT: Communi, Didier
; TITLE OF INVENTION: RECEPTOR GPCRxi10
; FILE REFERENCE: 9409/2082
; CURRENT APPLICATION NUMBER: US/09/885,453
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/885,453
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DNA sequence
; LOCATION: (1)..(1017)
; OTHER INFORMATION: GPCRxi6 DNA sequence
US-09-885-453-3

```

```

Query Match 93.0%; Score 1006; DB 10; Length 1017;
Best Local Similarity 99.9%; Pred. No. 2.8e-303;
Matches 1017; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 ATGCTGTCCATTGTTCTTCCAGGGGAGCAGAGCGGGAGCCGTCTGTGAGCTCTG 60
Db 1 ATGCTGTCCATTGTTCTTCCAGGGGAGCAGAGCGGGAGCCGTCTGTGAGCTCTG 60
Qy 61 CTCCTGTGAGGAGGCTCCCGGGACATGGAGAGTGGACATGAATACATCACAGGAACAA 120
Db 61 CTCCTGTGAGGAGGCTCCCGGGACATGGAGAGTGGACATGAATACATCACAGGAACAA 120
Qy 121 GGTCTCTGCCAGTCTCTCAGAGAACTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
Db 121 GGTCTCTGCCAGTCTCTCAGAGAACTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180
Qy 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGGCACTCTCTGGGCGCAAAACC 240
Db 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGGCACTCTCTGGGCGCAAAACC 240
Qy 241 AAGCGCTGGAGCTGTGCCACCACTCTAGTGTGTGCCACCACTGTGTTGCTGCCCC 300
Db 241 AAGCGCTGGAGCTGTGCCACCACTCTAGTGTGTGCCACCACTGTGTTGCTGCCCC 300
Qy 301 GTGCTATTGGCCCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 360
Db 301 GTGCTATTGGCCCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 360
Qy 361 CTGCTCTGCAAGCTGTGTCACCTTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
Db 361 CTGCTCTGCAAGCTGTGTCACCTTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
Qy 421 CTGACCTGCATCTCTGTGTCACCACTCTAGTGTGTGCCACCACTGTGTTGCTGCCCC 480
Db 421 CTGACCTGCATCTCTGTGTCACCACTCTAGTGTGTGCCACCACTGTGTTGCTGCCCC 480
Qy 481 TACCGGACCGCAGGAGCTGTGCTGGCAGCAGCACCCTGGGCGCTGTGTGCTCTC 540
Db 481 TACCGGACCGCAGGAGCTGTGCTGGCAGCAGCACCCTGGGCGCTGTGTGCTCTC 540
Qy 541 CAGCTGTGCTGCTGAGCTGTGCTGGCAGCTTCTCCACACGAGCTACATCAATGCGCAGATCATG 600
Db 541 CAGCTGTGCTGCTGAGCTGTGCTGGCAGCTTCTCCACACGAGCTACATCAATGCGCAGATCATG 600
Qy 601 TATGACATGACCACTGAGAGATTTTGTATCGGCTTTTGTCTACGGCATAGTCTGACA 660
Db 601 TATGACATGACCACTGAGAGATTTTGTATCGGCTTTTGTCTACGGCATAGTCTGACA 660

```


QY 661 TTCTGCGCTTTCTTCCCTCCTGTCATTTTGGTGTCTATTTCACTGATGGTCAGGAG 720
 Db 661 TTCTGCGCTTT-TTTCCCTCCTTGGTCACTTTTGGTGTCTATTTCACTGATGGTCAGGAG 719
 QY 721 CTTGATCAAGCCAGAGAGAACTCATGAGGACAGCAAGCCGAGGAGGTCAT 780
 Db 720 CTTGATCAAGCCAGAGAGAACTCATGAGGACAGCAAGCCGAGGAGGTCAT 779
 QY 781 CCGGACCATCTACTGCTGTGGCTCTTCAACCTCTCTTTTGGTGTCTTCAATATCAC 840
 Db 780 CCGGACCATCTACTGCTGTGGCTCTTCAACCTCTCTTTTGGTGTCTTCAATATCAC 839
 QY 841 TCGCTCTCTTACTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 900
 Db 840 TCGCTCTCTTACTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 899
 QY 901 AGCCAGTGTGGCTTACAGATATGAGGCTCTGTGGTGTGAGCAGCTGCTCAACCC 960
 Db 900 AGCCAGTGTGGCTTACAGATATGAGGCTCTGTGGTGTGAGCAGCTGCTCAACCC 959
 QY 961 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCAGAAACTGA 1018
 Db 960 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCAGAAACTGA 1017

RESULT 8

US-10-079-384-5
 ; Sequence 5, Application US/10079384
 ; Publication No. US20030108986A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Communi, Didier
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS COMPRISING G-PROTEIN COUPLED RECEPTORS
 ; FILE REFERENCE: 9409/2132
 ; CURRENT APPLICATION NUMBER: US/10/079,384
 ; CURRENT FILING DATE: 2002-02-20
 ; PRIOR APPLICATION NUMBER: US 09/885,453
 ; PRIOR FILING DATE: 2001-06-20
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 1017
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1017)
 ; OTHER INFORMATION:
 ; US-10-079-384-5

Query Match 93.0%; Score 1006; DB 15; Length 1017;
 Best Local Similarity 99.9%; Pred. No. 2.8e-303;
 Matches 1017; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 ATGCTGTCCATTTTGTCTTCTCCAGGGGAAGCAGAGCGGAGCGGTCTGTGAGCTCTG 60
 Db 1 ATGCTGTCCATTTTGTCTTCTCCAGGGGAAGCAGAGCGGAGCGGTCTGTGAGCTCTG 60
 QY 61 CTCTCGAGGAGGCTCCCGGACATGGAGAAGGTGGACATGAATACATCACAGGAACAA 120
 Db 61 CTCTCGAGGAGGCTCCCGGACATGGAGAAGGTGGACATGAATACATCACAGGAACAA 120
 QY 121 GGTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTTCTTCCCTGGCCTACAGTATC 180
 Db 121 GGTCTCTGCCAGTTCTCAGAGAGTACAAGCAAGTCTTCTTCCCTGGCCTACAGTATC 180
 QY 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGGCACTCTCTGGGCGCAAAACC 240
 Db 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTTGGCACTCTCTGGGCGCAAAACC 240
 QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCGGACCTGCTTTAT 300
 Db 241 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCGGACCTGCTTTAT 300

RESULT 9

US-10-023-586B-3
 ; Sequence 3, Application US/10023586B
 ; Publication No. US20030166882A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pfizer Ltd. (EP (GB) only)
 ; APPLICANT: Pfizer Inc. (US, JP, EP except GB)
 ; APPLICANT: Fidoock, Mark David
 ; TITLE OF INVENTION: No. US20030166882A1el Polypeptide
 ; FILE REFERENCE: PCI0960AGPR
 ; CURRENT APPLICATION NUMBER: US/10/023,586B
 ; CURRENT FILING DATE: 2001-12-18
 ; PRIOR APPLICATION NUMBER: GB 0030855.1
 ; PRIOR FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: US 60/260,563
 ; PRIOR FILING DATE: 2001-01-09
 ; PRIOR APPLICATION NUMBER: US 60/265,688
 ; PRIOR FILING DATE: 2001-02-01
 ; PRIOR APPLICATION NUMBER: GB 0101222.8
 ; PRIOR FILING DATE: 2001-01-17
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3

QY 301 GTGCTATTTGCCCTTCTCTCATCATCCTACTACTAGATGACAGGTGGCCCTTGGGGAG 360
 Db 301 GTGCTATTTGCCCTTCTCTCATCATCCTACTACTAGATGACAGGTGGCCCTTGGGGAG 360
 QY 361 CTGCTCTGCAAGCTGGTGCACCTTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
 Db 361 CTGCTCTGCAAGCTGGTGCACCTTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
 QY 421 CTGACCTGCAATCTGTGACACCACTTCTAGGTGTGGCACCACCTGTGTTCCGTGCCC 480
 Db 421 CTGACCTGCAATCTGTGACACCACTTCTAGGTGTGGCACCACCTGTGTTCCGTGCCC 480
 QY 481 TACCGGACCCGACGATGCTGCTGGGACACGACCACTCTGGGCTTGGTGGTCTCTC 540
 Db 481 TACCGGACCCGACGATGCTGCTGGGACACGACCACTCTGGGCTTGGTGGTCTCTC 540
 QY 541 CAGCTGTGCCCCACACTGGCTTCTCCACACGAGCTATCATCAATGCGCAGATGATCG 600
 Db 541 CAGCTGTGCCCCACACTGGCTTCTCCACACGAGCTATCATCAATGCGCAGATGATCG 600
 QY 601 TATGACATGACGACCAAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGACA 660
 Db 601 TATGACATGACGACCAAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGACA 660
 QY 661 TTGCTGGCTTTTCTTCCCTCCTTGGTCAATTTGGTGTGCTATTCTACTGATGGTCAGGAG 720
 Db 661 TTGCTGGCTTT-TTCCCTCCTTGGTCAATTTGGTGTGCTATTCTACTGATGGTCAGGAG 719
 QY 721 CTTGATCAAGCCAGAGAGAACTCATGAGGACAGCAAGCCGAGGAGGTCAT 780
 Db 720 CTTGATCAAGCCAGAGAGAACTCATGAGGACAGCAAGCCGAGGAGGTCAT 779
 QY 781 CCGGACCATCTACTGCTGTGGCTCTTCCACCTCTCTGTTTGTGGCTTCCATATCAC 840
 Db 780 CCGGACCATCTACTGCTGTGGCTCTTCCACCTCTCTGTTTGTGGCTTCCATATCAC 839
 QY 841 TCGCTCTCTTACTCACCATCTGCTTTCTTCTCAGGACTGCCAGCTCTTGATGGC 900
 Db 840 TCGCTCTCTTACTCACCATCTGCTTTCTTCTCAGGACTGCCAGCTCTTGATGGC 899
 QY 901 AGCCAGTGTGGCTTACAGATATGAGGCTCTGTGGTGTGAGCAGCTGCTCAACCC 960
 Db 900 AGCCAGTGTGGCTTACAGATATGAGGCTCTGTGGTGTGAGCAGCTGCTCAACCC 959
 QY 961 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCAGAAACTGA 1018
 Db 960 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCAGAAACTGA 1017


```

; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-586B-3

Query Match      92.7%; Score 1002.8; DB 15; Length 1020;
Best Local Similarity 99.6%; Pred. No. 2.8e-302;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1 ATGCTGTCATTTTCTTCTTCCAGGGGAGCAGAGCGGGAGCCCTGCTGAGCTCTG 60
DB 1 ATGCTGTCATTTTCTTCTTCCAGGGGAGCAGAGCGGGAGCCCTGCTGAGCTCTG 60
QY 61 CTCCTGGAGGAGGCTCCCGGGACATGGAGAGGTGGACATGAATACATACAGGAACAA 120
DB 61 CTCCTGGAGGAGGCTCCCGGGACATGGAGAGGTGGACATGAATACATACAGGAACAA 120
QY 121 GGTCTCTGCCAGTTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 180
DB 121 GGTCTCTGCCAGTTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 180
QY 181 ATCTTTATCTAGGCTGCGCACTAAATGGCACTCTTGTGGCACTCTCTGGGCGCAAAACC 240
DB 181 ATCTTTATCTAGGCTGCGCACTAAATGGCACTCTTGTGGCACTCTCTGGGCGCAAAACC 240
QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCGGACCTGCTTTAT 300
DB 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCGGACCTGCTTTAT 300
QY 301 GTGCTATGTCCTTCTCCTCATCATCACTACTACATGACAGTGGCCCTTCGGGGAG 360
DB 301 GTGCTATGTCCTTCTCCTCATCATCACTACTACATGACAGTGGCCCTTCGGGGAG 360
QY 361 CTGCTCTGCAAGCTGTGGCACTTCTCTTATATCAACCTTTACGCGCATCTCTGCTG 420
DB 361 CTGCTCTGCAAGCTGTGGCACTTCTCTTATATCAACCTTTACGCGCATCTCTGCTG 420
QY 421 CTGACCTGTCATCTCTGTCAGAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 480
DB 421 CTGACCTGTCATCTCTGTCAGAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 480
QY 481 ATCTTTATCTAGGCTGCGCACTAAATGGCACTCTTGTGGCACTCTCTGGGCGCAAAACC 540
DB 481 ATCTTTATCTAGGCTGCGCACTAAATGGCACTCTTGTGGCACTCTCTGGGCGCAAAACC 540
QY 541 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCGGACCTGCTTTAT 600
DB 541 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCGGACCTGCTTTAT 600
QY 601 TATCAGATGACCGAGGAGAACTTATGATCGGCTTTTGCCTACGCGCATGTTCTGACA 660
DB 601 TATCAGATGACCGAGGAGAACTTATGATCGGCTTTTGCCTACGCGCATGTTCTGACA 660
QY 661 TTGCTGCTGCTTTCTTTT - CCTCCTTGGTCAATTTTGTGTGCTATTCACTGATGTCAG 718
DB 661 TTGCTGCTGCTTTCTTTT - CCTCCTTGGTCAATTTTGTGTGCTATTCACTGATGTCAG 720
QY 719 AGCCTGATCAAGCCAGAGGAGAACTTATGATCGGCTTTTGCCTACGCGCATGTCGTC 778
DB 719 AGCCTGATCAAGCCAGAGGAGAACTTATGATCGGCTTTTGCCTACGCGCATGTCGTC 780
QY 779 ATCCGGACCACTCTACTGCTGTGGGCTCTTCAACCTCTGTTTGTGGCTTTCCATATC 838
DB 779 ATCCGGACCACTCTACTGCTGTGGGCTCTTCAACCTCTGTTTGTGGCTTTCCATATC 840
QY 839 ACTGCTCTCTTCTACTCAACATCTGCTTTCTGCTTTCTCAGGACGTCAGCTCTTGATG 898
DB 839 ACTGCTCTCTTCTACTCAACATCTGCTTTCTGCTTTCTCAGGACGTCAGCTCTTGATG 900
QY 899 GCAGCAGTGTGGCTACAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCTCTCAAC 958
DB 899 GCAGCAGTGTGGCTACAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCTCTCAAC 960
QY 901 GCACCCAGTGTGGCTACAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCTCTCAAC 960
DB 901 GCACCCAGTGTGGCTACAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCTCTCAAC 960

QY 959 CCAGTCTCTGCTACTTCTTTCAGGGGGGCAAAAATAGAGTCAAGGCTCTCTCAGAAACTGA 1018
DB 961 CCAGTCTCTGCTACTTCTTTCAGGGGGGCAAAAATAGAGTCAAGGCTCTCTCAGAAACTGA 1020

RESULT 10
US-10-763-972-3
; Sequence 3, Application US/10763972
; Publication No. US20040137500A1
; GENERAL INFORMATION:
; APPLICANT: PRIZER INC.
; APPLICANT: FIDOCK, Mark David
; TITLE OF INVENTION: Novel Polypeptide
; FILE REFERENCE: PC10960B
; CURRENT APPLICATION NUMBER: US/10/763,972
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: GB 0030855.1
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,563
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/265,688
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: GB 0101222.8
; PRIOR FILING DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-763-972-3

Query Match      92.7%; Score 1002.8; DB 17; Length 1020;
Best Local Similarity 99.6%; Pred. No. 2.8e-302;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1 ATGCTGTCATTTTCTTCTTCCAGGGGAGCAGAGCGGGAGCCCTGCTGAGCTCTG 60
DB 1 ATGCTGTCATTTTCTTCTTCCAGGGGAGCAGAGCGGGAGCCCTGCTGAGCTCTG 60
QY 61 CTCCTGGAGGAGGCTCCCGGGACATGGAGAGGTGGACATGAATACATACAGGAACAA 120
DB 61 CTCCTGGAGGAGGCTCCCGGGACATGGAGAGGTGGACATGAATACATACAGGAACAA 120
QY 121 GGTCTCTGCCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 180
DB 121 GGTCTCTGCCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 180
QY 181 ATCTTTATCTAGGCTGCGCACTAAATGGCACTCTTGTGGCACTCTCTGGGCGCAAAACC 240
DB 181 ATCTTTATCTAGGCTGCGCACTAAATGGCACTCTTGTGGCACTCTCTGGGCGCAAAACC 240
QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCGGACCTGCTTTAT 300
DB 241 AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCGGACCTGCTTTAT 300
QY 301 GTGCTATGTCCTTCTCCTCATCATCACTACTACATGACAGTGGCCCTTCGGGGAG 360
DB 301 GTGCTATGTCCTTCTCCTCATCATCACTACTACATGACAGTGGCCCTTCGGGGAG 360
QY 361 CTGCTCTGCAAGCTGTGGCACTTCTCTTATATCAACCTTTACGCGCATCTCTGCTG 420
DB 361 CTGCTCTGCAAGCTGTGGCACTTCTCTTATATCAACCTTTACGCGCATCTCTGCTG 420
QY 421 CTGACCTGTCATCTCTGTCAGAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 480
DB 421 CTGACCTGTCATCTCTGTCAGAGTACAGCAAGTCTACCTCTCCCTGGGCTACAGTATC 480
QY 481 TACCGGACCCGAGGAGGATGCTGGCTGGGCAACCGACACCTGGGCGCTGGTGGTCTC 540
DB 481 TACCGGACCCGAGGAGGATGCTGGCTGGGCAACCGACACCTGGGCGCTGGTGGTCTC 540
QY 541 CAGTGTGTCGCCACACTGGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTG 600
DB 541 CAGTGTGTCGCCACACTGGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTG 600
QY 601 TATCAGATGACCGAGGAGAACTTATGATCGGCTTTTGCCTACGCGCATGTTCTGACA 660
DB 601 TATCAGATGACCGAGGAGAACTTATGATCGGCTTTTGCCTACGCGCATGTTCTGACA 660
QY 661 TTGCTGCTGCTTTCTTTT - CCTCCTTGGTCAATTTTGTGTGCTATTCACTGATGTCAG 718
DB 661 TTGCTGCTGCTTTCTTTT - CCTCCTTGGTCAATTTTGTGTGCTATTCACTGATGTCAG 720
QY 719 AGCCTGATCAAGCCAGAGGAGAACTTATGATCGGCTTTTGCCTACGCGCATGTCGTC 778
DB 719 AGCCTGATCAAGCCAGAGGAGAACTTATGATCGGCTTTTGCCTACGCGCATGTCGTC 780
QY 779 ATCCGGACCACTCTACTGCTGTGGGCTCTTCAACCTCTGTTTGTGGCTTTCCATATC 838
DB 779 ATCCGGACCACTCTACTGCTGTGGGCTCTTCAACCTCTGTTTGTGGCTTTCCATATC 840
QY 839 ACTGCTCTCTTCTACTCAACATCTGCTTTCTGCTTTCTCAGGACGTCAGCTCTTGATG 898
DB 839 ACTGCTCTCTTCTACTCAACATCTGCTTTCTGCTTTCTCAGGACGTCAGCTCTTGATG 900
QY 899 GCAGCAGTGTGGCTACAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCTCTCAAC 958
DB 899 GCAGCAGTGTGGCTACAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCTCTCAAC 960
QY 901 GCACCCAGTGTGGCTACAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCTCTCAAC 960
DB 901 GCACCCAGTGTGGCTACAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCTCTCAAC 960
```

```

Db 541 CAGCTGCTGCCACACTGCGCTTCTCCACACGAGTACATCAATGGCCAGATGTCGG 600
QY 601 TATGACATCAGCAGCAGCAGATTTTGTATCGCTTTTGTGCTACGGATAGTCTTGACA 660
Db 601 TATGACATCAGCAGCAGCAGATTTTGTATCGCTTTTGTGCTACGGATAGTCTTGACA 660
QY 661 TTGTCTGGCTTTCTTTT-CCCTCCCTTGTGTCATTTTGGTGTGCTATTTCACTGATGTCAGG 718
Db 661 TTGTCTGGCTTTCTTTTCCCTCCCTTGTGTCATTTTGGTGTGCTATTTCACTGATGTCAGG 720
QY 719 AGCCTGATCAAGCCAGCAGCAGACCTCATGAGCAGAGCAACACAGCCCGAGCCAGTCC 778
Db 721 AGCCTGATCAAGCCAGCAGCAGACCTCATGAGCAGAGCAACACAGCCCGAGCCAGTCC 780
QY 779 ATCCGACCATCTACTGCTGTGTGCTCTTCAACCTCTGTTTGTGCTTCCATATC 838
Db 781 ATCCGACCATCTACTGCTGTGTGCTCTTCAACCTCTGTTTGTGCTTCCATATC 840
QY 839 ACTCGCTCTTCACTGCTGCTGCTTCTTCTGCTTCTCAGGATGCGCAGCTTGTATG 898
Db 841 ACTCGCTCTTCACTGCTGCTGCTTCTTCTGCTTCTCAGGATGCGCAGCTTGTATG 900
QY 899 GCAGCCAGTGTGCGCTTACAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 958
Db 901 GCAGCCAGTGTGCGCTTACAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 960
QY 959 CGAGTCTGTACTTCTTCAAGGGGGGCAAAAATAGATGATGCTGCTTCCAGAACTGA 1018
Db 961 CGAGTCTGTACTTCTTCAAGGGGGGCAAAAATAGATGATGCTGCTTCCAGAACTGA 1020

```

RESULT 11

```

US-10-333-946-20
; Sequence 20, Application US/10333946
; Publication No. US20040023252A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.
; APPLICANT: ARVIZU, Chandra S.; LAL, Preeti G.
; APPLICANT: BURFORD, Neil; YUE, Henry
; APPLICANT: GANDHI, Ameena R.; ELIOT, Vicki S.
; APPLICANT: RAMKUMAR, Jayalaxmi; BAUGHN, Mariah R.
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Nazinder K.
; APPLICANT: HAFALIA, April J.A.; YAO, Monique G.
; APPLICANT: LU, Yan; TRIBOULEY, Catherine M.
; APPLICANT: POLICKY, Jennifer L.; KEARNEY, Liam
; APPLICANT: GRAU, Richard C.; WARREN, Bridget A.
; APPLICANT: LEE, Ernestine A.; DING, Li
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0176 USN
; CURRENT APPLICATION NUMBER: US/10/333,946
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/23433
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/221,478
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/223,268
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/227,054
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 60/231,121
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,243
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 60/232,691
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/235,146
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 1076
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7474806CB1
US-10-333-946-20

Query Match          92.7%; Score 1002.8; DB 16; Length 1076;
Best Local Similarity 99.6%; Pred. No. 2.8e-302; Indels 2; Gaps 1;
Matches 1016; Conservative 0; Mismatches 2;

QY 1 ATGCTGTCCATTTTGTCTTCTTCCAGGGGAGCAGAGAGCGGAGCCCTCGTGGAGCTCTG 60
Db 57 ATGCTGTCCATTTTGTCTTCTTCCAGGGGAGCAGAGAGCGGAGCCCTCGTGGAGCTCTG 116
QY 61 CTCCTGAGGAGGAGCTCCCGGACATGGAGAGTGGACATGAATACATACAGAGAACAA 120
Db 117 CTCCTGAGGAGGAGCTCCCGGACATGGAGAGTGGACATGAATACATACAGAGAACAA 176
QY 121 GGTCTCTGCCAGTTCTCAGAGAACTACAAGCAAGTCTTACCTCTCCCTGGCTACAGTATC 180
Db 177 GGTCTCTGCCAGTTCTCAGAGAACTACAAGCAAGTCTTACCTCTCCCTGGCTACAGTATC 236
QY 181 ATCTTTATCTTAGGGCTGCCACCTAAATGGACATGCTCTTGTGGACCTCTCTGGGCGCAACC 240
Db 237 ATCTTTATCTTAGGGCTGCCACCTAAATGGACATGCTCTTGTGGACCTCTCTGGGCGCAACC 296
QY 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGGTGGCGGACCTCTCTTAT 300
Db 297 AAGCGTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGGTGGCGGACCTCTCTTAT 356
QY 301 GTGCTATTGCGCTTCTCTCATCATCACCTACTCTAGATGACAGGTGGCCCTTGGGGAG 360
Db 357 GTGCTATTGCGCTTCTCTCATCATCACCTACTCTAGATGACAGGTGGCCCTTGGGGAG 416
QY 361 CTGCTCTGCAAGCTGTGTCATCTCTGTTCTATATCAACCTTTACGCGAGCATCTGCTG 420
Db 417 CTGCTCTGCAAGCTGTGTCATCTCTGTTCTATATCAACCTTTACGCGAGCATCTGCTG 476
QY 421 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTGTGTGCGGCC 480
Db 477 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTGTGTGCGGCC 536
QY 481 TACCGGACCCGAGCAGTCTGCTGGGACACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 540
Db 537 TACCGGACCCGAGCAGTCTGCTGGGACACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 596
QY 541 CAGCTGTGCGCCACACTGGCCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG 600
Db 597 CAGCTGTGCGCCACACTGGCCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTGG 656
QY 601 TATGACATGACAGCAGCAGAGAAATTTGATCGGCTTTTGTGCTACGGCATAGTCTGACA 660
Db 657 TATGACATGACAGCAGCAGAGAAATTTGATCGGCTTTTGTGCTACGGCATAGTCTGACA 716
QY 661 TTGTCTGGCTTTCTTTT-CCCTCCCTTGTGTCATTTTGGTGTGCTATTTCACTGATGTCAGG 718
Db 717 TTGTCTGGCTTTCTTTTCCCTCCCTTGTGTCATTTTGGTGTGCTATTTCACTGATGTCAGG 776
QY 719 AGCCTGATCAAGCCAGCAGAGAGAACTCATAGAGCAGGCAACACAGCCCGAGCCAGTCC 778
Db 777 AGCCTGATCAAGCCAGCAGAGAGAACTCATAGAGCAGGCAACACAGCCCGAGCCAGTCC 836
QY 779 ATCCGACCATCTACTGCTGTGTGCTCTTCAACCTCTGTTTGTGCTTCCATATC 838
Db 837 ATCCGACCATCTACTGCTGTGTGCTCTTCAACCTCTGTTTGTGCTTCCATATC 896
QY 839 ACTCGCTCTTCACTGCTGCTGCTTCTTCTGCTTCTCAGGACTGCCAGCTCTTATG 898
Db 897 ACTCGCTCTTCACTGCTGCTGCTTCTTCTGCTTCTCAGGACTGCCAGCTCTTATG 956
QY 899 GCAGCCAGTGTGCGCTTACAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 958
Db 957 GCAGCCAGTGTGCGCTTACAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 1016

```

Qy	959	CCAGTCCTGTACTTTCTTTCAAGGGGCGCAAAATAGAGTCAAGCTCCTCCAGAACTGA	1018
Db	1017	CCAGTCCTGTACTTTCTTTCAAGGGGCGCAAAATAGAGTCAAGCTCCTCCAGAACTGA	1076

```

RESULT 12
US-10-088-726-25
; Sequence 25, Application US/10088726
; Publication No. US20030157558A1
; GENERAL INFORMATION:
; APPLICANT: Matsumoto et al.
; TITLE OF INVENTION: NOVEL GUANOSINE TRIPHOSPHATE-BINDING PROTEIN-COUPLED RECEPTORS AND
; TITLE OF INVENTION: THEREOF, AND PRODUCTION AND USES THEREOF
; FILE REFERENCE: 62514
; CURRENT APPLICATION NUMBER: US/10/088,726
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: PCT/JP00/09408
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: JP 1999-375152
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: JP 2000-101339
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-088-726-25

```

Db	541	TTTGATCGGCTTTTGGCTACGGGCAATAGTCTGACATGTCTGGCTTCTTTCCCTCCTT	600
Qy	685	GGTCATTTTGGTGTGCTATTCACTGATGGTCAGGAGCCTGATCAAGCCAGAGGAACT	744
Db	601	GGTCATTTTGGTGTGCTATTCACTGATGGTCAGGAGCCTGATCAAGCCAGAGGAACT	660
Qy	745	CATCAGACAGGCAACACAGCCCGAGCCAGGTCCATCGGNACCATCTACTGGTGTGG	804
Db	661	CATCAGACAGGCAACACAGCCCGAGCCAGGTCCATCGGACCATCTACTGGTGTGG	720
Qy	805	CCTCTTCAACCTCTGTTTGTGGCCCTTCCATATCACTCGCTCCTTCTACCTCACCATCTG	864
Db	721	CCCTTCAACCTCTGTTTGTGGCCCTTCCATATCACTCGCTCCTTCTACCTCACCATCTG	780
Qy	865	CTTTCTCTTTCTCAGGACTGCCAGCTCTTGTATGGCAGCCAGTGTGGCTTCAAGATATG	924
Db	781	CTTTCTCTTTCTCAGGACTGCCAGCTCTTGTATGGCAGCCAGTGTGGCTTCAAGATATG	840
Qy	925	GAGGCTCTGGTGTGTGACGAGCTGCCTCAACCCAGTCTGTACCTTTCTTTCAAGGG	984
Db	841	GAGGCTCTGGTGTGTGACGAGCTGCCTCAACCCAGTCTGTACCTTTCTTTCAAGGG	900
Qy	985	GGCAAAATATAGATCAGGCTCCTCCAGAAACTGAGGCAGAACAGTTGGGTGACATCCA	1044
Db	901	GGCAAAATATAGATCAGGCTCCTCCAGAAACTGAGGCAGAACAGTTGGGTGACATCCA	960
Qy	1045	GCTGGGAGGAAGAGATGCCCAGGGTTGAAACAGATCTGG	1082
Db	961	GCTGGGAGGAAGAGATGCCCAGGGTTGAAACAGATCTGG	998

```

RESULT 13
US-10-275-910-6
; Sequence 6, Application US/10275910
; Publication No. US20030166142A1
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Shyam
; TITLE OF INVENTION: REGULATION OF HUMAN P2Y-LIKE G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 4974.00885
; CURRENT APPLICATION NUMBER: US/10/275,910
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/203,582
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/269,857
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 6
; LENGTH: 851
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-275-910-6

```

Query Match	92.1%	Score 996.4	DB 15	Length 1002
Best Local Similarity	99.9%	Pred. No. 2.7e-300		
Matches 997	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	85	ATGGAGAAGGTGGACATGAATACATACAGAAACAAGGTCTCTGCCAGTTCCTCAGAGAAG	144	
Db	1	ATGGAGAAGGTGGACATGAATACATACAGAAACAAGGTCTCTGCCAGTTCCTCAGAGAAG	50	
QY	145	TACAAGCAAGTCTACCTCTCCCTGGCCCTA CAGTATCACTTTTATCTTAGGGCTGCACCTA	204	
Db	61	TACAAGCAAGTCTACCTCTCCCTGGCCCTA CAGTATCACTTTATCTTAGGGCTGCACCTA	120	
QY	205	AATGSCACTGCTTGTGTGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACACC	264	
Db	121	AATGSCACTGCTTGTGTGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACACC	180	
QY	265	TATCTGTGAACCTGATGTGGCCGACCTGCTTTATGTGCTATTGCCCTTCTCTATCATCTC	324	
Db	181	TATCTGTGAACCTGATGTGGCCGACCTGCTTTATGTGCTATTGCCCTTCTCTATCATCTC	240	
QY	325	ACCTACTCACTAGATGACAGTGTGCCCTCTCGGGAGCTGCTCTGCAAGCTGTGCACATTC	384	
Db	241	ACCTACTCACTAGATGACAGTGTGCCCTCTCGGGAGCTGCTCTGCAAGCTGTGCACATTC	300	
QY	385	CTGTTCTATATCAACCTTTACGGCAGACATCTGCTGTGACCTGTGATCTCTGTGCACCAG	444	
Db	301	CTGTTCTATATCAACCTTTACGGCAGACATCTGCTGTGACCTGTGATCTCTGTGCACCAG	360	
QY	445	TTCTTAGTGTGTGCCACCCACTGTGTTTGGCTGCCTCTACGGACCCGAGGACATGCTGG	504	
Db	361	TTCTTAGTGTGTGCCACCCACTGTGTTTGGCTGCCTCTACGGACCCGAGGACATGCTGG	420	
QY	505	CTGGCACCAGCACCACTCTGGGCCCTGGTGGTCTCTCAGCTGTGCCACACTGGCCCTTC	564	
Db	421	CTGGCACCAGCACCACTCTGGGCCCTGGTGGTCTCTCAGCTGTGCCACACTGGCCCTTC	480	
QY	565	TCCACACGGACTACATCAATGGCCAGATGATCTGTTATGACATGACCCAGCCAGAGAAAT	624	
Db	481	TCCACACGGACTACATCAATGGCCAGATGATCTGTTATGACATGACCCAGCCAGAGAAAT	540	
QY	625	TTTGATCGGCTTTTGTGGCTACGGCAAGTGTCTGACATTGTCTGGCTTTCTTTCCCTCCTT	684	


```

; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-974C-13

```

```

Query Match      47.1%; Score 510; DB 10; Length 510;
Best Local Similarity 100.0%; Pred. No. 1.8e-148;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 TGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCGACCTGCTTTATGTGCTA 306
Db 1 TGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCGACCTGCTTTATGTGCTA 60

Qy 307 TTGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGAGCTGCTC 366
Db 61 TTGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGAGCTGCTC 120

Qy 367 TGAAGCTGTGACCTTCCTGTTCTATATCAACCTTTACGGAGCATCTCTGCTGAC 426
Db 121 TGAAGCTGTGACCTTCCTGTTCTATATCAACCTTTACGGAGCATCTCTGCTGAC 180

Qy 427 TGATCTCTGTGACCACTTCCTAGGTGTGTGCCACCACTGTGTTGCTGCCCTACCG 486
Db 181 TGATCTCTGTGACCACTTCCTAGGTGTGTGCCACCACTGTGTTGCTGCCCTACCG 240

Qy 487 ACCGCAAGGATGCTGGCTGGGCAACAGCACCTGGGCCCTGTGTGTCCTCCAGCTG 546
Db 241 ACCGCAAGGATGCTGGCTGGGCAACAGCACCTGGGCCCTGTGTGTCCTCCAGCTG 300

Qy 547 CTGCCACACTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGGTATGAC 606
Db 301 CTGCCACACTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGGTATGAC 360

Qy 607 ATGACCAAGCAAGAGATTTTGATCGGCTTTTTCCTACGGCATAGTCTTGACATTTGCT 666
Db 361 ATGACCAAGCAAGAGATTTTGATCGGCTTTTTCCTACGGCATAGTCTTGACATTTGCT 420

Qy 667 GGCTTTCTTCTCCCTCTGCTGCTATTTTGGTGTGCTATTACATGATGGTCAGGAGCTGAT 726
Db 421 GGCTTTCTTCTCCCTCTGCTGCTATTTTGGTGTGCTATTACATGATGGTCAGGAGCTGAT 480

Qy 727 CAAGCCAGAGGAGAACCTCATGAGGACAGG 756
Db 481 CAAGCCAGAGGAGAACCTCATGAGGACAGG 510

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 08:44:17 ; Search time 114 Seconds
(without alignments)
6746.259 Million cell updates/sec

Title: US-10-763-972-1
Perfect score: 1082
Sequence: 1 atgctgccattttgttcc.....ccagggttgacacatctgg 1082

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	227.8	21.1	2025	4	US-09-016-434-1482
2	227.8	21.1	2025	4	US-09-016-434-1482
3	203.8	18.8	1842	1	US-08-442-134A-1
4	203.8	18.8	1842	1	US-08-444-581B-1
5	203.8	18.8	1842	1	US-08-446-088A-1
6	178.6	16.5	1429	4	US-08-016-434-1068
7	178.6	16.5	1429	4	US-09-077-173D-1
8	164.8	15.2	984	3	US-08-513-974B-57
9	164.8	15.2	984	4	US-09-461-436B-57
10	164.8	15.2	1023	3	US-08-513-974B-379
11	163.2	15.1	1571	4	US-09-016-434-1108
12	156.8	14.5	984	3	US-08-459-046-1
13	156.8	14.5	984	4	US-08-102-170B-1
14	148.8	13.8	984	3	US-08-513-974B-41
15	148.8	13.8	984	4	US-09-461-436B-41
16	148.8	13.8	1020	3	US-08-513-974B-370
17	97.4	9.0	1773	4	US-09-016-434-1405
18	95.8	8.9	998	4	US-08-432-174A-3
19	95.6	8.8	1892	4	US-09-900-699A-1
20	92	8.5	1475	1	US-08-097-938-1
21	92	8.5	1475	1	US-08-476-000-1
22	92	8.5	1475	1	US-08-472-840-1
23	92	8.5	1475	2	US-08-476-976-1
24	92	8.5	1475	3	US-08-474-410-1
25	92	8.5	1475	3	US-08-486-673B-1
26	92	8.5	2732	1	US-08-478-000-60
27	92	8.5	2732	1	US-08-472-840-60

28	92	8.5	2732	2	US-08-476-976-60
29	92	8.5	2732	3	US-08-474-410-60
30	92	8.5	2732	3	US-08-486-673B-60
31	91.4	8.4	1205	1	US-08-417-103-13
32	91.4	8.4	1634	1	US-07-816-283-1
33	91.4	8.4	1634	1	US-08-417-103-1
34	91.4	8.4	1634	4	US-09-016-434-1302
35	89.8	8.3	3055	4	US-03-016-434-1456
36	88.4	8.2	1020	4	US-09-170-496D-31
37	88.4	8.2	1020	4	US-09-170-496D-181
38	88.4	8.2	1900	4	US-09-016-434-1484
39	88.4	8.2	1901	1	US-08-153-848-43
40	88.4	8.2	1901	3	US-09-299-843A-43
41	88.4	8.2	1901	3	US-09-088-337B-43
42	88.4	8.2	1901	5	PCT-US93-11153-43
43	88.4	8.2	2453	5	PCT-US95-07180-1
44	87.4	8.1	1119	4	US-09-170-496D-65
45	87.4	8.1	1119	4	US-09-170-496D-199

ALIGNMENTS

RESULT 1
US-09-016-434-1482
; Sequence 1482, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSES: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1482:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2025 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G984506
; US-09-016-434-1482

Query Match 21.1%; Score 227.8; DB 4; Length 2025;
Best Local Similarity 56.4%; Pred. No. 6.8e-53;

Matches 486; Conservative 0; Mismatches 367; Indels 8; Gaps 3;
 QY 127 TGCAGATTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTT 186
 Db 318 TGCAGATTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTT 377
 QY 187 ATCTAGAGGCTGCCACTAAATGGCACTGTCTTTGGCACTCTCTGGGCGCAACCAAGCGC 246
 Db 378 GTGCTTGGGCTGTCTGAAGCGGCTGTCTACATCTTCTTGTGCGGCTCAAGACC 437
 QY 247 TGGAGCTGTGCACCACTATCTGGTGAACCTGTATGTTGGCGGACCTGTCTTATGTGCTA 306
 Db 438 TGGAAATGCGTCCACCAATATATGTTCCACCTGTGTCTGATGCACTGTATGCGGCC 497
 QY 307 T---TGCCCTTCTCATCATCACTACTACTAGTATGAGTGGCCCTTCGGGAGCTG 363
 Db 498 TCCCTGCGCTGCTGTCTATTAATCAAGCGGCGGCGGCGGCTGTGCTGCGGCTG 557
 QY 364 CTCTGCAAGCTGTGCACTTCTCTGTATATCAACTTTACGGCAGATCTCTGCTGCTG 423
 Db 558 CTCTGCAAGCTGTGCGGCTTCTCTTACACCAACTTTTACTGCAAGTCTCTTCTCCTC 617
 QY 424 ACCTGCACTCTGTGCACTTCTCTAGTGTGCGCACTCTGCTGCGGCGGCTGTGCTGCTG 483
 Db 618 ACCTGCACTCTGTGCACTTCTCTAGTGTGCGCACTCTGCTGCGGCGGCTGTGCTGCTG 677
 QY 484 CGGACCCGCGAGCATGCTTGGCTGGCACTGAGTGGCCCTTCGGGAGCTGCTTCAATG 543
 Db 678 GCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
 QY 544 CTGCTGCGCACTGCGGCTTCTCCCACTGAGTGTGCGCACTGAGTGTGCTGCTGCTGCTG 603
 Db 738 GCGACCATCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 797
 QY 604 GACATGACCGAGAGAGATTTTGTGCTTACGGCATAGTCTTCTGACATG 663
 Db 798 GACACTCGGCACTGCGGCTTCTCAGCGCTTCTGCGCTTACAGTCACTGCTGCGG 857
 QY 664 TCTGGCTT--TCTTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 Db 858 CTGCTCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917
 QY 722 CTGATCAAGCGAGAGAGTCTCATGAGGACAGGCAACACAGCCGCGGCGGCTGAGTGG 781
 Db 918 CTGTAAGCGAGAGTCTCATGAGGACAGGCAACACAGCCGCGGCGGCTGAGTGG 977
 QY 782 CGGACCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 841
 Db 978 GCGACCATCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037
 QY 842 CGCTCTTCTACCTCACTGCTTCTGCTTCTCAGGAGTGGCGGCTTCTGATGGA 901
 Db 1038 CGCACCTCTACTACTCTCTCCGCT---CGTGGACCTCAGCTGCCACACCTCAAGCC 1094
 QY 902 GCCAGTGTGGCTACAAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCTCAACCCA 961
 Db 1095 ATCAACATGGCTACAAGTTACCGGCGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTG 1154
 QY 962 GTCTGTACTTCTTCAAG 982
 Db 1155 GTGCTACTTCTGCTGGTGG 1175

RESULT 2

US-09-814-915A-74
 ; Sequence 74, Application US/09814915A
 ; Patent No. 6750015
 ; GENERAL INFORMATION:
 ; APPLICANT: Horwitz, Kathryn
 ; APPLICANT: Richer, Jennifer
 ; TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relat
 ; TITLE OF INVENTION: Thereto
 ; FILE REFERENCE: 2848-39

; CURRENT APPLICATION NUMBER: US/09/814,915A
 ; CURRENT FILING DATE: 2002-03-21
 ; PRIOR APPLICATION NUMBER: 60/214,870
 ; PRIOR FILING DATE: 2000-06-28
 ; NUMBER OF SEQ ID NOS: 108
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 74
 ; LENGTH: 2025
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-814-915A-74
 Query Match 21.1%; Score 227.8; DB 4; Length 2025;
 Best Local Similarity 56.4%; Pred. No. 6.8e-53;
 Matches 486; Conservative 0; Mismatches 367; Indels 8; Gaps 3;
 QY 127 TGCAGATTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTT 186
 Db 318 TGCAGATTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTT 377
 QY 187 ATCTAGAGGCTGCCACTAAATGGCACTGTCTTTGGCACTCTCTGGGCGCAACCAAGCGC 246
 Db 378 GTGCTTGGGCTGTCTGAAGCGGCTGTCTACATCTTCTTGTGCGGCTCAAGACC 437
 QY 247 TGGAGCTGTGCACCACTATCTGGTGAACCTGTATGTTGGCGGACCTGTCTTATGTGCTA 306
 Db 438 TGGAAATGCGTCCACCAATATATGTTCCACCTGTGTCTGATGCACTGTATGCGGCC 497
 QY 307 T---TGCCCTTCTCATCATCACTACTACTAGTATGAGTGGCCCTTCGGGAGCTG 363
 Db 498 TCCCTGCGCTGCTGTCTATTAATCAAGCGGCGGCGGCGGCTGTGCTGCGGCTG 557
 QY 364 CTCTGCAAGCTGTGCACTTCTCTGTATATCAACTTTACGGCAGATCTCTGCTGCTG 423
 Db 558 CTCTGCAAGCTGTGCGGCTTCTCTTACACCAACTTTTACTGCAAGTCTCTTCTCCTC 617
 QY 424 ACCTGCACTCTGTGCACTTCTCTAGTGTGCGCACTCTGCTGCGGCGGCTGTGCTGCTG 483
 Db 618 ACCTGCACTCTGTGCACTTCTCTAGTGTGCGCACTCTGCTGCGGCGGCTGTGCTGCTG 677
 QY 484 CGGACCCGCGAGCATGCTTGGCTGGCACTGAGTGGCCCTTCGGGAGCTGCTTCAATG 543
 Db 678 GCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
 QY 544 CTGCTGCGCACTGCGGCTTCTCCCACTGAGTGTGCGCACTGAGTGTGCTGCTGCTGCTG 603
 Db 738 GCGACCATCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 797
 QY 604 GACATGACCGAGAGAGATTTTGTGCTTACGGCATAGTCTTCTGACATG 663
 Db 798 GACACTCGGCACTGCGGCTTCTCAGCGCTTCTGCGCTTACAGTCACTGCTGCGG 857
 QY 664 TCTGGCTT--TCTTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 Db 858 CTGCTCTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 917
 QY 722 CTGATCAAGCGAGAGAGTCTCATGAGGACAGGCAACACAGCCGCGGCGGCTGAGTGG 781
 Db 918 CTGTAAGCGAGAGTCTCATGAGGACAGGCAACACAGCCGCGGCGGCTGAGTGG 977
 QY 782 CGGACCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 841
 Db 978 GCGACCATCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037
 QY 842 CGCTCTTCTACCTCACTGCTTCTGCTTCTCAGGAGTGGCGGCTTCTGATGGA 901
 Db 1038 CGCACCTCTACTACTCTCTCCGCT---CGTGGACCTCAGCTGCCACACCTCAAGCC 1094
 QY 902 GCCAGTGTGGCTACAAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCTCAACCCA 961
 Db 1095 ATCAACATGGCTACAAGTTACCGGCGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTG 1154
 QY 962 GTCTGTACTTCTTCAAG 982

309	TCCTCGCGCTGCTGGTCTATTACTACGCCCGCGGCGACCACTGGCCCTTCAGACAGGTG	360
364	CTCTGCAAGCTGGTGACATTCCTGCTTATATCAACTTTAOCGACGATCTGCTGCTG	423
369	CTCTGCAAGCTGGTGCGCTTCTCTTCTACACCAACTTTTACTGACGATCTCTTCTCTC	428
424	ACCTGCATCTCTGTGSCACAGTTCCTAGGTGTGTGCCACCACTGTGTTCGCTGCCTAC	483
429	ACTTGCATCAGCGTGCACCGGTGTCTGGGCGTCTTACGACCTCTGGCTCCCTGGCTGG	488
484	CGGACCCCGCAGGCATTCCTGCTGGGCAACGACGACCACTTGGGCCCTTGTGTCTCCAG	543
489	GGCGCGGCGCGCTACGCTCGCGGTGGCGGGCGCGGTGTGGG---TGTTGTGTGCTGGCC	545
544	CTGCTGCCCACTACGCTTCTCCCAACGAGCTATACATCAATGSCCAGATGATCTGGTAT	603
546	TGCGAGGGCCCGTCTCTACTTTGTCAACACAGCGCGCGCGGCGCTTACCTTGCAC	605
604	GACATGACCCAGCAGAGAAATTTTGATCGGCTTTTGCCCTACGGCATAGTTCTGACATG	663
606	GACACTCGGCACCGAGCTCTTCAGCGCTTTCGTTGGCCCTACAGCTCAGTCATGCTGGGC	665
664	CTGTGGCTT---TCCTTTCCTCTTGGTCAATTTTGGTGTGCTATTCACTGATGCTCAGGAGC	721
666	CTGCTCTTCGCGGTGCCCTTTTCCGCTCATCTTGTCTGTTCATCGTGTCTATGCTCGGCGA	725
722	CTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGGTCCATC	781
726	CTGTAAAGCCAGCCTACGGGACCTCTGGGCGGCTGCTAGGGCCAAAGCGCAAGTCCGTC	785
782	CGGACATCCTACTGGTGTGGCTCTTCAACCTCTGTGTTTGTGGCTTCCATATCACT	841
786	CGCACCATCGCGTGTGCTGTCTTCCGCTCTGTCTCTGCAATTTCCACGTCACC	845
842	CGCTCCTTCACTCACCATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGCA	901
846	CGCACCTCTACTACTCTCTTCTGCT---CGCTGACCTCAGCTGCCACACCTCAGCGCC	902
902	GCCAGTGTGCGCTACAAGATATGAGGCGCTCTGCTGATGTGAGCAGCTGCCTCAACCCA	961
903	ATCAACATCGCCTACAAGT---TACCGGCTGCCAGTGTAACTTGCCTTGACCCG	959
962	GTCCCTGACTTTCTTCAAGG	982
960	GTGCTCTACTTCTCGCTGGG	980

RESULT 4
 US-08-444-581B-1
 ; Sequence 1, Application US/08444581B
 ; Patent No. 5607836
 ; GENERAL INFORMATION:
 ; APPLICANT: Boucher, Richard C.
 ; APPLICANT: Weisman, Gary A.
 ; APPLICANT: Turner, John T.
 ; APPLICANT: Harden, Thomas K.
 ; APPLICANT: Parr, Claude E.
 ; APPLICANT: Sullivan, Daniel M.
 ; APPLICANT: Erb, Laura
 ; APPLICANT: Lustig, Kevin D.
 ; TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
 ; TITLE OF INVENTION: Null Cells Expressing P2U Receptors
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Bell, Seltzer, Park & Gibson
 ; STREET: Post Office Drawer 34009
 ; CITY: Charlotte
 ; STATE: No. 5607836th Carolina
 ; COUNTRY: USA
 ; ZIP: 28234
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,581B
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/442,134
FILING DATE: 16-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 5470-71A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-420-2200
TELEFAX: 919-881-3175
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1842 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 57..1181
US-08-444-581B-1

Query Match 18.8%; Score 203.8; DB 1; Length 1842;
Best Local Similarity 56.3%; Pred. No. 2.7e-46;
Matches 485; Conservative 0; Mismatches 362; Indels 14; Gaps 5;
QY 127 TGGCAGTTCTCAGAGAGTCAAGCAAGTCTACTCTCCCTGGCTACATATCATCTTT 186
DB 129 TGGCGCTTCAACGAGGACTTCAAGTACGTGTGTCTGTCTGTCTACGCGGTGTGC 188
QY 187 ATCTAGGCTGCCAATAATGGACATCTTGTGGACATCTGGGCGCAACCAAGCGC 246
DB 189 GTGCTTGGGCTGTCTGAAGCGGTGGGCTCTAATCTTGTGGCGCTCAAGACC 248
QY 247 TGGAGCTGTGCCACCATCTCTGTGGTGAACCTGTATGTGGCGGACCTGCTTTATGTCTA 306
DB 249 TGGAAATGGCTCCACCATATATGTTCCACTGTGTGTCTGTATGACATCTGTATGGGCC 308
QY 307 T---TGGCTTCTCATCATCATCTACTACTAGATGACAGGTGGCGCTTGGGGAGCTG 363
DB 309 TCCCTGGCGTGGTGTCTATTACTAGCGCGGCGGACCACTGGCGCTTTCAGACGGTG 368
QY 364 CTCTGCAAGCTGTGTGACCTTCTCTTATATCAACCTTTACGGCAGCATCTCTGCTGCTG 423
DB 369 CTCTGCAAGCTGTGTGCGCTTCTCTTCTACCAACCTTTTACTGACGATCTCTTCTTC 428
QY 424 ACCTGATCTCTGTGACAGGATCTCTAGGTGTGTGCCACCATCTGTCTGCTGGCGCTAC 483
DB 429 ACCTGATCATGCGTGCACCGGTGTCTGCGGCTCTTACGACCTCTGCGCTCTCCCTGCG 488
QY 484 CGGACCGCAGGATGTCTGTGGGACCAAGACCACTGGGCGCTTGTGGTCTCTCCAG 543
DB 489 GGCGGGCGGCTACGCTGCGCGGCTGGCGGGCGGTGTGGG---TGTGTGTGTGGCC 545
QY 544 CTGTGCGCCACACTGGCGCTTCTGCCACAGCATATCAATGCGCAGATGATCTGGTAT 603
DB 546 TGGCAGGCGCGGTGTCTATTGTGTACCAAGCGCGCGCGCTTAACTGCCAC 605
QY 604 GACATGACAGCAAGAGATTTGTGCGCTTTTGGCTACGCGATGTTCTGACATG 663
DB 606 GACACTGGGACCCGAGCTCTTACGCGCTTGTGGCGCTTACAGCTCAGTCTGTGGCG 665
QY 664 TCTGGCTT---TCTTTCCCTCTTGGTCAATTTGGTGTGCTATTCATGTATGGTTCAGGAGC 721
DB 666 CTGCTCTTGGGCGGCTTGGCGGTCTATCTTGTCTGTGTACGTGTCTATGGCTCGGCGA 725
QY 722 CTGATCAAGCAGGAGACCTTCATGAGCAGGACCAACAGACCGCGGCGGCTCCATC 781

DB 726 CTGTAAAGCCAGCTACGGGACCTCGGGCGGCGCTCGGCGCAAGCGCAAGTCCGTG 785
QY 782 CGGACATCTACTGTGTGTGGCTCTTCACTCTCTTGTGCGCTTGTGCGCTTCCATATCAT 841
DB 786 CGCACCATCGCGGT 845
QY 842 CGCTCTCTTACTCTACCTCACCATCTGTCTTCTCTAGGACTGCCAGCTCTTGTATGSCA 901
DB 846 CGCACCCTTACTACTCTCTCTCGCT---CGCTGGACCTCAGTCCACACCTCAAGGCC 902
QY 902 GCAGTGTGGCTCAAGATATGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 961
DB 903 ATCAACATGGGCTACAGGT---TACCGGCTGGCGCAGTGTAAACAGTTGCTTGAACCC 959
QY 962 GTCCTGTACTTCTTCTCAAGG 982
DB 960 GTGCTTACTTCTTCTGGTGGG 980

RESULT 5

US-08-446-088A-1
Sequence 1, Application US/08446088A
Patent No. 5691156

GENERAL INFORMATION:

APPLICANT: Boucher, Richard C.
APPLICANT: Weisman, Gary A.
APPLICANT: Turner, John T.
APPLICANT: Harden, Thomas K.
APPLICANT: Parr, Claude E.
APPLICANT: Sullivan, Daniel M.
APPLICANT: Erb, Laura
APPLICANT: Lustig, Kevin D.

TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
TITLE OF INVENTION: Null Cells Expressing P2U Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell, Seltzer, Park & Gibson
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: No. 5691156th Carolina
COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/446,088A

FILING DATE: 19-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kenneth D. Sibley

REGISTRATION NUMBER: 31,665

REFERENCE/DOCKET NUMBER: 5470-71C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-420-2200

TELEFAX: 919-881-3175

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1842 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 57..1181

US-08-446-088A-1

Query Match 18.8%; Score 203.8; DB 1; Length 1842;
Best Local Similarity 56.3%; Pred. No. 2.7e-46;

	Matches	485; Conservative	0; Mismatches	362; Indels	14; Gaps	5
Qy	127	TGCCAGTTCTCAGAGAGTACAAGCAAGTTACTCTTCCCTGGCCCTACAGTATCATCTTT	186			
Db	129	TGCGCTTCAACAGGAGCTTCAAGTAGTGCTGCTGCTGTCTCTACGCGCTGGTGTGC	188			
Qy	187	ATCCTAGGGCTGCCACTAAATGGCACTGTCTTGTGGACATCTCTGGGGCCAAACCAAGCGC	246			
Db	189	GTGCTTGGGCTGTGTTGAACGCGGTGGGCTCTTACATCTTCTGTGTGGCGCTCAAGACC	248			
Qy	247	TGGAGCTGTGCCACCACCTATCTGGTGAACCTGATGTGGCCGACCTGCTTATGTGCTTA	306			
Db	249	TGGAATGTGCTCCACACATATATGTTCCACCTGGCTGTCTCATCTGATGCTATGCGCC	308			
Qy	307	T---TGGCCTTCTCATCATCACTACTCAGTAGATGACAGGTGGCGCTTCGGGAGCTG	363			
Db	309	TCCCTGCGCTGCTGGTCTATTACTAGCGCGGGGACCACTGGCGCTTCAAGACGGTG	368			
Qy	364	CTCTGAAGCTGGTGCATCTCTGTTCTATATCAACCTTTACGGCAGCATCTGTGCTGTG	423			
Db	369	CTCTGCAAGCTGGTGGCTTCTCTTCTACACCAACCTTTACTGCAGCATCTCTTCCTC	428			
Qy	424	ACCTGCATCTGTGTGCACAGTTCTTAGTGTGTGTGCCACCACTGTGTTCGTGCGCTAC	483			
Db	429	ACCTGCATCAGGGTCACCGGTGTCTGGGGGTCTTACAGCTCTGTGGCTCCCTGCGCTGG	488			
Qy	484	CGSACCGCAGGCATGCTGTGGCTGGCACCAAGCACCACTGGGGCCCTGTGTGTTCTCCAG	543			
Db	489	GGCGGGGCCGTACGCTCGCGGGTGGCGGGCCGTGTGGG---TGTGTGTGTGGCC	545			
Qy	544	CTGTGCGCACACTGGCCTTTCTCCCAACGGACTACATCAATGGCCAGATGATGTGTAT	603			
Db	546	TGCCAGGCCCGTGTCTCTACTTTGTCAACCACGCGCGGGGCGCTAACTGCCAC	605			
Qy	604	GACATGACAGCCAGAGAAATTTGATCGGCTTTTGCTTACGGCATGTTCTGACATG	663			
Db	606	GACACTGGGCACCGAGCTCTTACGCGCTTGTGTGGCTTACAGCTCATGTCATGCTGGGC	665			
Qy	664	TCTGGCTT---TCTTTTCCCTCTTGCTCATTTTGGTGTGCTATTTCATGTATGTCAGGAGC	721			
Db	666	CTGCTCTTCGGCGTGCCTTTGGCGTCACTCTTGTCTGTTACGTGCTCATGGCTGGCGA	725			
Qy	722	CTGATCAAGCCAGAGAGAACTCATGAGNAGGCGAAACACACCCGAGCGCAGGTCCATC	781			
Db	726	CTGTAAAGCCAGCTACCGGACCTCGGGGCGGCTGCTCTAGGGCCCAAGCGCAAGTCCG	785			
Qy	782	CGGACCATCTACTGGTGTGTGGCTCTTCAACCTCTGTGTTTTGTGGCTTCCATATCACT	841			
Db	786	CGCACCAATCGCGTGGTGTGCTTCTGCGCTCTGTCTTCTGGCATTCACGTCACC	845			
Qy	842	CGTCTCTTCTACTCATCCATCTGTTTTCTGTGTTTTCTCAGGATCTGCAGGTCTTGATGGCA	901			
Db	846	CGCACCTCTACTACTCTCTTCCGT---CGGTGGACCTCAGCTGCCACACCTCTCAACGCC	902			
Qy	902	GCCAGTGTGGCCTACAAGATATGAGGCGCTCTGTGTGATGTGAGCAGCTGCTCAACCCA	961			
Db	903	ATCACATGGCCTACAAGT---TACCGGCTGCCAGTGTCTAACAGTTGGCTTGAACCC	959			
Qy	962	GTCTGTACTTTTTCGAAG	982			
Db	960	GTGCTCTACTTCTGGCTGGG	980			

RESULT 6
US-09-016-434-1068
; Sequence 1068, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Jasice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490

CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 59/016,434
FILING DATE: HEREWITH

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1068:
SEQUENCE CHARACTERISTICS:
LENGTH: 1429 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1124904
9-016-434-1068

Query Match	16.5%;	Score 178.6;	DB 4;	Length 1429;
Best Local Similarity	54.08;	Pred. No. 2.1e-394;	Indels 16;	Gaps 5;
Matches 481;	Conservative 0;	Mismatches 39;		
Qy	126	CTGCCAGTTCTCAGAGAAGTACAGCAAGTCTACCTCTCTCCCTGCGCTACAGTATCATCTTT	185	
Db	258	CTGTGGTTTGATGAGGATTTCAAGTTTCATCTCTGCTGCCTGTGAGCTATGCAGTTGTCTTT	317	
Qy	186	TATCTTAGGCTCGCCATAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCG	245	
Db	318	TGTGCTGGGTTTGGGCTTTAAAGCCGCCAACCCCTATGGCTCTTTCATCTTCGCTCCGACCTCGGACC	377	
Qy	246	CTGGAGCTGTGCCACCACCACTTATCTGTTGAACCTCATGTGGTGGCGGACCTGCTTTATGTGCT	305	
Db	378	CTGGGATGCAACGGCCACCCTACATGTTTCCACCTGGCATTTTCAGACACCTTGTATGTGCT	437	
Qy	306	AT---TGCCCTTCTCATCATCACTACTCACTAGATGACAGGTGGCGCTTCGGGAGCT	362	
Db	438	GTGCGTGGCCACCCTCATCTACTATTATGAGGCCAACCAACCACTGGGCGCTTTTGGCACTGA	497	
Qy	363	GCTCTGCAAGCTGGTGGACATCTCCTGTCTCTATATCAACCTTTTAGCGAGCATCTCTGCTGCT	422	
Db	498	GAICTGCAAGTTGTCGCGCTTTCTTTTCTATTGGAACTCTACTGCAGTGTCCCTTTTCCCT	557	
Qy	423	GACCTGGATCTCTGTGGACCAAGTTCTTAGTGTGTGGCCACCACCTGTTGTGGCTGCCCTTA	482	
Db	558	CAGCTGCATCAGGCGTGCACCGCTACCTGGGCACTCTGCCACCCACTTTCGGGCACCTACGCTG	617	
Qy	483	CCGGACCCGACGACATCGCTGGCTGGGCACACGACCACTCGGCGCCCTGGTGTGCTCCCA	542	
Db	618	GGGCGCGCTCTCGCTCGAGCGCTTCTCTGCTCGGCAAGTTTGGTTGTCGTAGCCGCGCTG	677	
Qy	543	GCTGCTGCCCACTGTGGCTTTTCCACACGGAATCATCATCAATGGCCAGATGATCTGGTA	602	
Db	678	CCTGCTGCCCAACCTGTTCTTTTGTTCACAAACAGCAACAAAGGGACCAACCGTCTGTGCCA	737	

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/513,974
FILING DATE: 14-SEP-1995
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
APPLICATION NUMBER: 7-093989
FILING DATE: 19-APR-1995
APPLICATION NUMBER: 7-057186
FILING DATE: 16-MAR-1995
APPLICATION NUMBER: 7-007177
FILING DATE: 20-JAN-1995
APPLICATION NUMBER: 6-326611
FILING DATE: 28-DEC-1994
APPLICATION NUMBER: 6-270017
FILING DATE: 02-NOV-1994
APPLICATION NUMBER: 6-236357
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: 6-236356
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: 6-189274
FILING DATE: 11-AUG-1994
APPLICATION NUMBER: 6-189273
FILING DATE: 11-AUG-1994
APPLICATION NUMBER: 6-189272
FILING DATE: 11-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: CONLIN, DAVID G.
REGISTRATION NUMBER: <Unknown>

REFERENCE/DOCKET NUMBER: 45753 DIV2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-439-4444
TELEFAX: 617-439-4170

INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:

LENGTH: 984 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 57:

US-09-461-436B-57

Query Match 15.2%; Score 164.8; DB 4; Length 984;

Best Local Similarity 52.0%; Pred. No. 1.1e-35;

Matches 444; Conservative 0; Mismatches 402; Indels 8; Gaps 3;

```
QY 126 CTGCCAGTCTCAGAGAAGTACAGCAAGTCTACTCTCCCTGGCCCTACAGTATCATCTT 185
DB 51 CTGTGTCTACCGCGAGAAGCTTCAAGCAACTGCTGCTGCCACTGTGTATTGGCGGTGCT 110
QY 186 TATCTAGGCTGCCACTAAATGACACTGCTTGTGGCACTCTCTGGGGCCAAACCAAGCG 245
DB 111 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 170
QY 246 CTGAGCTGTGCCACCACTATCTGTGTAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
DB 171 CTTGACCGCGACGGCGGTGACACCCCTAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 230
QY 306 AT---TGCCCTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTCGGGAGCT 362
DB 231 CTCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 290
QY 363 GCTCTGAAGTGTGTGCACTTCTCTTCTATATCAAGCTTTACGGCAGCATCTCTGCTGCT 422
DB 291 CGCCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 350
QY 423 GACCTGATCTCTGTGACCAAGTCTCTAGGTGTGTGCAACCACTGTTGGTGGCTTA 482
DB 351 CACCTGATCAAGTCTTCCAGCGCTACCTGGGGATCTGCCACCGCTGGCCCTTGGGACAA 410
QY 483 CC---GGACCGCGAGGATGCTGCTGGGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 539
DB 411 ACGTGGGGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 470
```

RESULT 10

US-08-513-974B-379
; Sequence 379, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ontaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ogi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-057186
; FILING DATE: 16-MAR-1995

```
QY 540 CCAGCTGTGCTGCCACACTGGCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTG 599
DB 471 CCAGTGTGCTGCCACAGCCATCTTCGCTGCCACAGGATCCAGCGTAAACGCACTGTCTG 530
QY 600 GTATGATGACCAAGCAAGAGATTTTGATGGGCTTTTGGCTACGCGATAGTCTGAC 659
DB 531 CTATGACCTCAGCGCGCTGCTGCCCTGGCCACCACTATATGCCCCTATGGCATGGCTCTCAC 590
QY 660 ATTGCTGTGCTTCT - TTCCCTCTCTGCTCAATTTGGTGTGCTATTCACTGATGCTCAG 717
DB 591 TGTATCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650
QY 718 GAGCTGTATCAAGCAGAGGAGAACCTCATGAGGACAGGACAGCAGCCGAGCCAGGTC 777
DB 651 CGCCCTGTGCGCCAGGATGGCCCGCAGAGCTGTGGCCAGGAGCGCGCTGGCAAGGC 710
QY 778 CATCGGACCATCTTACTGCTGTGGCTCTTCAACCTCTGTTTGTGGCTTCCATAT 837
DB 711 GCGCCGATGCGCGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
QY 838 CACTGCTCTTCTTACTCATCATCTGCTTCTGCTTCTGCTTCTCAGGACTGCCAGCTTGTAT 897
DB 771 CACCAAGACAGCTTACTTGGCAGTGGGCTCGACGCGCGCGCTGCCCTGCACTGTATTGGA 830
QY 898 GCGAGCAGTGTGGCTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCCTCAA 957
DB 831 GGCCTTTCAGCGGCTTCAAGGACGCGCGCTTCCAGTGCACACAGCGTGTGCTGGA 890
QY 958 CCCAGTCTCTGACT 971
DB 891 CCCCATCTCTCT 904
```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 379:
SEQUENCE CHARACTERISTICS:
LENGTH: 1023 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 37..1020
US-08-513-974B-379

Query Match 15.28; Score 164.8; DB 3; Length 1023;
Best Local Similarity 52.08; Pred. No. 1.2e-35;
Matches 444; Conservative 0; Mismatches 402; Indels 8; Gaps 3;

QY 126 CTGCCAGTTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTT 185
DB 87 CTGTGTCTACGGCAGAACTTCAAGCAACTGCTGCTGCCACCTGTGTATTTCGGCGGTCT 146
QY 186 TATCCTAGGGTGCACATAATGCACTGTCTTGTGCACTCTCTGGGCCAAACCAAGCG 245
DB 147 GCGCGCTGGCTGCGCTGCAACATCTGTGTATACCAAGATCTGCACGTCCCGCGCGGC 206
QY 246 CTGAGCTGTGCCACCACTATCTGTGTGAACCTGTAGTGGCGCAGCTCTGTATGTGCT 305
DB 207 CTTGACCGCAGCGCGGTGACACCTTAACCTTGTCTGTGCTGACCTGTATATGCTG 266
QY 306 AT---TGCCTTCTCATCATCTACCTACTCATAGATGACAGGTGGCCCTTCGGGGAGCT 362
DB 267 CTCCTGCGCCCTGCTCATCTACAACTATGCCAAGTGTGATCACTGGCCCTTTGGCGACT 326
QY 363 GCTCTGCAAGCTGTGCACTTCTGTCTTATATCAACCTTTACGGCAGCATCTCTGTGCT 422
DB 327 CGCTGCGCGCTGTGCTGCTTCTTCTTCTATGCAACCTGTGACGGCAGCATCTCTTCT 386
QY 423 GACCTGATCTCTGTGACCAAGTTCCTAGGTGTGTGCAACCACTGTGTTTCGTGCGCTA 482
DB 387 CACCTGATCAGCTTCCAGCGCTACCTGGGATCTGCAACCGCTGGCCCTGGGCAAA 446
QY 483 CC---GGACCGGAGGATGCTGGCTGGGCAACCACTGGCCCTGTGGTCTCT 539

DB 447 ACGTGGGGCGCGCGGCTGCTGCTAGTGTGTAAACCGTGTGGCTGGCGGTGACAAC 506
QY 540 CCAGCTGCTGCCACACTGGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTG 599
DB 507 CCAGTGGCTGCCACAGCCATCTTGGCTGCCACAGGATCCAGCGTAACCCGACTGTCTG 566
QY 600 GTATGACATGACAGCCAGAGAAATTTTGTATCGGCTTTTCCCTACGGCATAGTCTGAC 659
DB 567 CTATGACCTCAGCCCGCTGCGCTGGCCACCACTATATGCCCCTATGGCATGGCTCTCAC 626
QY 660 ATTGCTGCGCTTCT--TTCCCTCCTTGGTCAATTTTGGTGTGCTATTCACTGATGGTCA 717
DB 627 TGTATGCGCTTCTGCTGCGCTTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 686
QY 718 GAGCCTGATCAAGCCAGAGAGAACCTCATGAGACAGGCAACACAGCCCGAGCGAGTCT 777
DB 687 CGGCTGTGCCCGAGGATGCCCGGAGAGCTGTGGCCAGGAGCGCGCTGGCAAGGC 746
QY 778 CATCGGACCATCTACTGCTGTGGCTCTTCCACCTCTGTTTGTGCCCTTCCATAT 837
DB 747 GCGCGCATGCGCGGTGGTGGCTGTGCGCTTTGCCATAGCTTCTGCTGCTTTTCACT 806
QY 838 CACTGCTCTTCTACCTCAACATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGAT 897
DB 807 CACCAAGACAGCTACCTGCGAGTGGGCTCGACGCCGGCGCTCCCTGCACTGTATTGA 866
QY 898 GCGACCGAGTGGCTTACAAGATATGAGAGCTCTGAGGCTCTGAGTGTGAGCAGCTGCTCAA 957
DB 867 GCGCTTTGACGGGCTTACAAAGCAGCGCGCTTTGCCAGTCCCAACAGCGTCTGGA 926
QY 958 CCAGCTCTCTACT 971
DB 927 CCGCATCTCTCT 940

RESULT 11
US-09-016-434-1108
Sequence 1108, Application US/09016434
Patent No. 6500936
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1108:

SEQUENCE CHARACTERISTICS:
LENGTH: 1571 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g1296659
US-09-016-434-1108

Query Match 15.1%; Score 163.2; DB 4; Length 1571;
Best Local Similarity 51.9%; Pred. No. 3.9e-35;
Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;

```
QY 126 CTGCAGTTCACAGAGTACAGAGTCTACCTCTCCCTGGCCCTACAGTATCATCTT 185
DB 126 CTGCAGTTCACAGAGTACAGAGTCTACCTCTCCCTGGCCCTACAGTATCATCTT 185
QY 327 CTGTGCTACCGCGAGAACTTCAAGCACTGCTGCTGCCACTGTGTATTCGGCGGTCT 386
DB 327 CTGTGCTACCGCGAGAACTTCAAGCACTGCTGCTGCCACTGTGTATTCGGCGGTCT 386
QY 186 TATCTAGGCTGCCACTAAATGGCACTGTCTTGGCACTCTCTGGGGCCAAACCAAGCG 245
DB 186 TATCTAGGCTGCCACTAAATGGCACTGTCTTGGCACTCTCTGGGGCCAAACCAAGCG 245
QY 397 GCGGCTGGCTGCGCTGGAACATCTGTGTCAITACCAATCTGACGTCGCCCGCGGC 446
DB 397 GCGGCTGGCTGCGCTGGAACATCTGTGTCAITACCAATCTGACGTCGCCCGCGGC 446
QY 246 CTGAGGCTGTGCCACCACTATCTGTGAACTGATGCTGGCCGACCTGCTTTATGTCT 305
DB 246 CTGAGGCTGTGCCACCACTATCTGTGAACTGATGCTGGCCGACCTGCTTTATGTCT 305
QY 447 CTTGACCGCGACGCGGTGTACACCTTAACCTTGTCTGCTGACCTGTATATGCTG 506
DB 447 CTTGACCGCGACGCGGTGTACACCTTAACCTTGTCTGCTGACCTGTATATGCTG 506
QY 306 AT---TGCCCTTCTCATCATCACTACTCACTAGATGACAGTGGCCCTTCGGGAGCT 362
DB 306 AT---TGCCCTTCTCATCATCACTACTCACTAGATGACAGTGGCCCTTCGGGAGCT 362
QY 507 CTCCTGCCCCGTCTATCACTAACAATATGCCAAGTGATCACTGGGCCCTTTGGCGACT 566
DB 507 CTCCTGCCCCGTCTATCACTAACAATATGCCAAGTGATCACTGGGCCCTTTGGCGACT 566
QY 363 GCTCTGAAGCTGTGCACTTCTCTTCTATATCAACTTACGGCAGCATCTGCTGCT 422
DB 363 GCTCTGAAGCTGTGCACTTCTCTTCTATATCAACTTACGGCAGCATCTGCTGCT 422
QY 567 CGCTGCGCTGTGCTGCTTCTTCTATGCAACTGCAAGGAGATCTCTTCTCT 626
DB 567 CGCTGCGCTGTGCTGCTTCTTCTATGCAACTGCAAGGAGATCTCTTCTCT 626
QY 423 GACCTGATCTGTGACCAAGTCTTAGGTGTGTGCAACCACTGTTGCTGGCCTA 482
DB 423 GACCTGATCTGTGACCAAGTCTTAGGTGTGTGCAACCACTGTTGCTGGCCTA 482
QY 627 CACCTGATCACTTCCAGCGCTACCTGGGCACTGCGCCGCTGGCCCTCGGCACAA 686
DB 627 CACCTGATCACTTCCAGCGCTACCTGGGCACTGCGCCGCTGGCCCTCGGCACAA 686
QY 483 CC---GGACCGCAGGATGCTGCTGGCGGCAACCACTGCGCCCTGCTGCTCT 539
DB 483 CC---GGACCGCAGGATGCTGCTGGCGGCAACCACTGCGCCCTGCTGCTCT 539
QY 687 ACCTGGGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
DB 687 ACCTGGGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 746
QY 540 CCAGCTGCTGCCACACTGCGCTTCTCCCAACGCACTACATCAATGGCCAGATCTG 599
DB 540 CCAGCTGCTGCCACACTGCGCTTCTCCCAACGCACTACATCAATGGCCAGATCTG 599
QY 747 CCAGTGTGCTGCCACAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
DB 747 CCAGTGTGCTGCCACAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 806
QY 600 GTATGACATGACCAAGCAAGATTTTGTATCGCTTTTGGCTACGGCATGCTGAC 659
DB 600 GTATGACATGACCAAGCAAGATTTTGTATCGCTTTTGGCTACGGCATGCTGAC 659
QY 807 CTATGACCTCAGCCCGCTGCGCTGCGCCACCCACTATATGCCCTATGGCATGCTCTAC 866
DB 807 CTATGACCTCAGCCCGCTGCGCTGCGCCACCCACTATATGCCCTATGGCATGCTCTAC 866
QY 660 ATTGCTGGCTTCT---TTCCCTCTTGTGCTATTTTGTGCTGCTTCACTGATGCTG 717
DB 660 ATTGCTGGCTTCT---TTCCCTCTTGTGCTATTTTGTGCTGCTTCACTGATGCTG 717
QY 867 TGTATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
DB 867 TGTATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
QY 718 GAGCTGATCAAGCAGAGGAGAACCTATGAGAGACAGGCAACACAGCCCGACCCAGGTC 777
DB 718 GAGCTGATCAAGCAGAGGAGAACCTATGAGAGACAGGCAACACAGCCCGACCCAGGTC 777
QY 927 CCGCTTGTGCGCGCAGATGCGCCGCGAGAGCTGTGGCCCGAGGCGCGCTGGCAAGGC 986
DB 927 CCGCTTGTGCGCGCAGATGCGCCGCGAGAGCTGTGGCCCGAGGCGCGCTGGCAAGGC 986
QY 778 CATCCGACCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
DB 778 CATCCGACCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
QY 987 GCGCGCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
DB 987 GCGCGCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
QY 838 CACTGCTCTTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 897
DB 838 CACTGCTCTTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 897
QY 1047 CACCAAGACAGCTTACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
DB 1047 CACCAAGACAGCTTACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
QY 898 GGCAGCAGTGTGCTTACAGATATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
DB 898 GGCAGCAGTGTGCTTACAGATATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
QY 1107 GGCCTTTGAGCGGCTTACAAAGGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166
DB 1107 GGCCTTTGAGCGGCTTACAAAGGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1166
QY 958 CCCAGTCTCTTACT 971
DB 958 CCCAGTCTCTTACT 971
QY 1167 CCCATCTCTTCT 1180
DB 1167 CCCATCTCTTCT 1180
```

RESULT 12

US-08-459-046-1
Sequence 1. Application US/08459046
Patent No. 608039
GENERAL INFORMATION:

APPLICANT: Coleman, Roger
APPLICANT: Au-Young, Janice
APPLICANT: Stuart, Susan G.
TITLE OF INVENTION: A No. 608039el Human Purinergic P2U Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCITE PHARMACEUTICALS, INC.
STREET: 3330 Hillview Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,046

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PF-0038 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: Placenta
CLONE: 179696

US-08-459-046-1

Query Match 14.5%; Score 156.8; DB 3; Length 984;

Best Local Similarity 51.4%; Pred. No. 1.8e-33;

Matches 439; Conservative 0; Mismatches 407; Indels 8; Gaps 3;

```
QY 126 CTGCAGTTCACAGAGTACAGAGTCTACCTCTCCCTGGCCCTACAGTATCATCTT 185
DB 126 CTGCAGTTCACAGAGTACAGAGTCTACCTCTCCCTGGCCCTACAGTATCATCTT 185
QY 51 CTGTGCTACCGCGAGAACTTCAAGCACTGCTGCTCCCACTGTGTATTCGGCGGTCT 110
DB 51 CTGTGCTACCGCGAGAACTTCAAGCACTGCTGCTCCCACTGTGTATTCGGCGGTCT 110
QY 186 TATCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCG 245
DB 186 TATCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCG 245
QY 111 GCGCGCTGCCCTCCCGTGAACATCTGTGTCAITACCAAGTATCTGCAGTCTCCCGCGGC 170
DB 111 GCGCGCTGCCCTCCCGTGAACATCTGTGTCAITACCAAGTATCTGCAGTCTCCCGCGGC 170
QY 246 CTGAGGCTGTGCCACCACTATCTGGTGAACCTGATGCTGGCCGACCTGCTTTATGTCT 305
DB 246 CTGAGGCTGTGCCACCACTATCTGGTGAACCTGATGCTGGCCGACCTGCTTTATGTCT 305
QY 171 CTTGACCGCGAGCGCGTGTACACCTTAACCTTGTCTGCTGCTGCTGCTGCTGCTGCTG 230
DB 171 CTTGACCGCGAGCGCGTGTACACCTTAACCTTGTCTGCTGCTGCTGCTGCTGCTGCTG 230
QY 306 AT---TGCCCTTCTCATCATCACTACTCACTAGATGACAGTGGCCCTTCGGGAGCT 362
DB 306 AT---TGCCCTTCTCATCATCACTACTCACTAGATGACAGTGGCCCTTCGGGAGCT 362
QY 231 CTCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 290
DB 231 CTCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 290
QY 363 GCTCTGAGGCTGTGCACTTCTCTTCTATATCAACTTACCGCAGCATCTCTGCTGCT 422
DB 363 GCTCTGAGGCTGTGCACTTCTCTTCTATATCAACTTACCGCAGCATCTCTGCTGCT 422
QY 291 GCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 350
DB 291 GCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 350
QY 423 GACCTGCACTCTCTGTGCAACCAAGTTCCTAGGTGTGTGCGCACCACTGTGCTGCTGCT 482
DB 423 GACCTGCACTCTCTGTGCAACCAAGTTCCTAGGTGTGTGCGCACCACTGTGCTGCTGCT 482
```

Db 351 CACCTGCATCAGCTTCACGCGTACCTGGGCACTGCGACCCGCTGGCCCTTGGCAAA 410
Qy 483 CC---GGACCCGAGGATACCTGGCTGGGCAACGACACACCTGGCCCTGGTGTCT 539
Db 411 ACCTGGGGGCGCGGGCTGCTGGCTAGTGTGTAGCGCTGGCTGGCCGCTGCAAC 470
Qy 540 CCAGCTGCTGCCACACAGCTGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTG 599
Db 471 CCAGTGCCTGCCACAGCCTCTTGGCTGCCACAGGACTCCAGGTAACCGCACTGCTG 530
Qy 600 GTATGACATACACAGGCAAGAGAAATTTGTAGTGGCTTTTGGCTACCGGATGATCTGAC 659
Db 531 TTATGACCTCAGCCCGCTGCTGGCCACCCACTATATGCCCTATGGGATGGCTCTCAC 590
Qy 660 ATTGTCTGGCTTCT---TTCCCTCTCTGGCTGCTATTTGGTGTGCTATTTCACTGATGTG 717
Db 591 TGTATCGGCTTCTGCTGCTGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650
Qy 718 GAGCTGATCAAGCAGGAGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCCAGGTC 777
Db 651 CCGCTGTCGCGCAGGATGCGCGGCGAGAGCCTGTGGCCAGAGGCGGCTGGCAAGGC 710
Qy 778 CATCGGACCATCTACTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
Db 711 GGCCGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 770
Qy 838 CACTCGCTCTTCTACCTCACCCTGCTGCTTCTGCTTCTCAGGACTGCGCAGCTCTTGAT 897
Db 771 CACCAAGACGCTTACTGGCAGTGCCTCGACGCGCGGCTGCTGCTGCTGCTGCTGCTG 830
Qy 898 GGCAGCAGTGTGGCTTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAA 957
Db 831 GGCCTTTGCGCGGCTTACAAAGCACGCGCGCTTTTSCAGTGCCTGCAACAGCGTGTGA 890
Qy 958 CCCAGTCTGTACT 971
Db 891 CCCATCTCTTCT 904

RESULT 13

US-09-102-710B-1
; Sequence 1, Application US/09102710B
; Patent No. 6479630
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger
; APPLICANT: Au-Young, Janice
; APPLICANT: Stuart, Susan G.
; TITLE OF INVENTION: A NOVEL HUMAN PURINERGIC P2U RECEPTOR
; FILE REFERENCE: PP-0038-1 DIV
; CURRENT APPLICATION NUMBER: US/09/102,710B
; CURRENT FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6479630 179696CB1
US-09-102-710B-1

Query Match 14.5%; Score 156.8; DB 4; Length 984;
Best Local Similarity 51.4%; Pred. No. 1.8e-33;
Matches 439; Conservative 0; Mismatches 407; Indels 8; Gaps 3;
Qy 126 CTGCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTT 185
Db 51 CTGTGCTACCGGAGACTTCAAGCACTGCTGCTCCACCTGTGATTTCGGGGTGTCT 110
Qy 186 TATCTAGGCTGCCACTTAATGGCACTGTCTTGTGGCACTCTCTGGGCGCCAAACAGCG 245
Db 111 GCGCCTGCCCTCCGCTGAACATCTGTGTCTATTACCCAGATCTGCACTCCCGCGGGC 170

Qy 246 CTGGAGCTGTGCCACCACTATCTGTTGAACTGATGTTGGCCGACCTGCTTTATGTCT 305
Db 171 CTTGACCCGCGAGGCGGTGTACACCTTAAACCTTGTCTGCTGCTGCTGCTATATGCTG 230
Qy 306 AT---TGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTGGGGAGCT 362
Db 231 CTCCCTGCCCCCTGCTCATCTACACTATGCCCCAAGTGTATCACTGGCCCTTTGGCGACT 290
Qy 363 GCTCTGCAAGCTGGGCACTTCTCTTCTATATCAACCTTTACGCGACATCTCTGCTCT 422
Db 291 CGCCTGCGCCCTGGTTCCTCTTCTATGCCAACCTGTCACGGGAGATCTCTTCTCT 350
Qy 423 GACCTGCACTCTCTGTGTCACCACTTCTTAGTGTGTGTCACCACTGTGTTGCTGCCCTA 482
Db 351 CACCTGCACTAGCTTCCAGCGCTACTTGGCATCTGCCACCGCTGGCCCTGCGCAAA 410
Qy 483 CC---GGACCCCGCAGGATGCTGCTGGGCAACAGCACCACTGGGCGCTTGGTCTCT 539
Db 411 AGTGGGGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 470
Qy 540 CCAGCTGTGCGCCACACTGGCCCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTG 599
Db 471 CCAGTGCCTGCGCCACAGCCTCTTCTGCTGCCACAGGCTCCAGCGTAACCGCACTGCTG 530
Qy 600 GTATGACATGACCGCAAGAGAAATTTGTATCGGCTTTTGGCTTACGGGATAGTCTGAC 659
Db 531 TTATGACCTCAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 590
Qy 660 ATTGTCTGGCTTCT---TTCCCTCTCTGCTCATTTTGGTGTGCTATTCACTGATGCTG 717
Db 591 TGTATCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650
Qy 718 GAGCTGTATCAAGCCAGAGGAGAACCTCATAGGACAGGCAACACAGCCCGAGCGAGTCT 777
Db 651 CCGCTGTGCGCCAGGATGCGCGGCGAGAGCTGTGCGCCAGAGCGCGCTGGCAAGGC 710
Qy 778 CATCGGACCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
Db 711 GCGCCGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 770
Qy 838 CACTGCTCTCTTCTACCTCACCCTGCTGCTTCTGCTTCTCAGGACTGCGCAGCTCTTGAT 897
Db 771 CACCAAGACGCTTACTGGCAGTGCCTCGACGCGCGGCTGCTGCTGCTGCTGCTGCTG 830
Qy 898 GGCAGCAGTGTGGCTTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAA 957
Db 831 GGCCTTTGCGCGGCTTACAAAGCACGCGCGCTTTTSCAGTGCCTGCAACAGCGTGTGA 890
Qy 958 CCCAGTCTGTACT 971
Db 891 CCCATCTCTTCT 904

RESULT 14

US-08-513-974B-41
; Sequence 41, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ontaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ogi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA

COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513.974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-513-974B-41

Query Match 13.8%; Score 148.8; DB 3; Length 984;
Best Local Similarity 50.8%; Pred. No. 2.9e-31;
Matches 434; Conservative 0; Mismatches 412; Indels 8; Gaps 3;
QY 126 CTGCAGTCTCAGAGAGTCAAGCAAGTCTACCTCTCCCTGCGCTACAGTATCATCTTT 185
DB 51 CTTCGCTCTACCGCTGAGGATTTCAAGCGACTGCTGCTAAACCCCGGTATCTCGGTGTGCT 110
QY 186 TATCCTAGGCTGCGCACTTAATGCACTCTGCTGCACTCTCTGGCGCAACACCGAGCG 245
DB 111 GGTGGTGGCTGCGCACTGAACATCTCGGTCAATTGCGCACTCGCGCACTCCCGCGGAC 170
QY 246 CTGAGCTGTGCCACCACTATCTGTGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCT 305

DB 171 CTTGACCCGTTCCGCTGTGTACACCTGAACTGGCAGCTGGCGGACCTGATGTATGCTG 230
QY 306 AT---TGCCCTTCTCATCATCATCTACTACTAGATGACAGGTGGCCCTTCGGGAGCT 362
DB 231 TTCACTACCCCTACTTATCTATACTACGCCAGAGGGGACCACTGGCCCTTCGGAGACT 290
QY 363 GCTCTGCAAGCTGGTGCACTTCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCT 422
DB 291 CGCTGCGCTTTGTACGCTTCTCTTCTATGCCAATCTACATGGCAGCATCTCTTCT 350
QY 423 GACCTGCACTCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTGTGTGCTCCCTA 482
DB 351 CACCTGCACTTACCTTCCAGCGCTACTGGGCACTGTGCCACCCCTGGCTTCTTGGCACA 410
QY 483 CC---GGACCCGAGGCACTGCTGCTGGGCAACAGCACTGCTGGCCCTGGTGTGCT 539
DB 411 GCGTGGAGGTGCGCTGCTGCTTGGGTAGTGTGTGAGTCTGTGTGCTGGCTGGTACAGC 470
QY 540 CCAGCTGTGCCACACTGGCCCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTG 599
DB 471 CCAGTGCCTGCCACGCGAGTCTTTGCTGCCACAGGCACTCCAGCGCAACCGCACTGTG 530
QY 600 GTATGACATGACAGCCAGCAAGAAATTTGATGCGGCTTTTGGCTACGGCAGTATTTCTG 659
DB 531 CTACGACCTGAGCCCACTCTCTACTGCTTACTGCTTACTGCTTCTGCTTCTGCTTCTG 590
QY 660 ATTGTCTGCT---TTCTTTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
DB 591 GGTGATGCGCTTCTTGTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTG 650
QY 718 GAGCCTGATCAAGCCAGAGGAGAACTCTATGAGGACAGGCAACACAGCCCGAGCGCTC 777
DB 651 CCGCCTGTGTCGCCAGGATGGCCAGAGGCTCTGTGCCCAAGAGCGCGCAGCAAGGC 710
QY 778 CATCGGACCATCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
DB 711 GGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
QY 838 CACTGCTCTCTTCTACTCACCATCTGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTG 897
DB 771 CACCAAGACAGCTACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830
QY 898 GCGACCGAGTGGGCTACAGATATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
DB 831 GACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 890
QY 958 CCCAGTCTGCTACT 971
DB 891 CCCCATCTCTTCT 904

RESULT 15

US-09-461-436B-41
; Sequence 41, Application US/09461436B
; Patent No. 6538107
; GENERAL INFORMATION:
; APPLICANT: Shuji Hinuma
; Ryo Fujii
; Yasuaki Ito
; TITLE OF INVENTION: G Protein Coupled Receptor Protein,
; Production, And Use Thereof
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edwards & Angell, LLP
; STREET: 101 Federal Street
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02209
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/461,436B
FILING DATE: 14-Dec-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/513,974
FILING DATE: 14-SEP-1995
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
APPLICATION NUMBER: 7-093989
FILING DATE: 19-APR-1995
APPLICATION NUMBER: 7-057186
FILING DATE: 16-MAR-1995
APPLICATION NUMBER: 7-007177
FILING DATE: 20-JAN-1995
APPLICATION NUMBER: 6-326611
FILING DATE: 28-DEC-1994
APPLICATION NUMBER: 6-270017
FILING DATE: 02-NOV-1994
APPLICATION NUMBER: 6-236357
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: 6-236356
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: 6-189274
FILING DATE: 11-AUG-1994
APPLICATION NUMBER: 6-189273
FILING DATE: 11-AUG-1994
APPLICATION NUMBER: 6-189272
FILING DATE: 11-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: CONLIN, DAVID G.

REGISTRATION NUMBER: <Unknown>

REFERENCE/DOCKET NUMBER: 45753 DIV2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-439-4444

TELEFAX: 617-439-4170

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 984 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 41:

US-09-461-436B-41

Query Match 13.8%; Score 148.8; DB 4; Length 984;
Best Local Similarity 50.8%; Pred. No. 2.9e-31;
Matches 434; Conservative 0; Mismatches 412; Indels 8; Gaps 3;

```
Qy 126 CTGCCAGTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTT 185
Db 51 CTGGCTGTACCGTGAGGATTTCAAGGACCTGCTCTAACCCCGGTATCTCGGTGGTCT 110
Qy 186 TATCTAGGGTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCG 245
Db 111 GGTGTGGCTGGCTGCCACTGAACATCTGGTCATTGGCCAGATCTGGCATCCCGCCGGAC 170
Qy 246 CTGAGCTGTGCCACCACTATCTGTGAACCTGATGGTGGCCGACCTGTTTATGTGCT 305
Db 171 CCTGACCCGTTCCGCTGTGTACACCTGAACCTGGCACTGGCCGACCTGATGTATGCTG 230
Qy 306 AT--TGCCCTTCCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCT 362
Db 231 TTCCTACCCCTACTTATCTATCTAGTACAGGAGGGGACCACTGGCCCTTCGGAGACCT 290
Qy 363 GCTCTGCAAGCTGGTGCACTTCTCTGTCTTATATCAACTTTACGGCAGCATCTCTGCTGCT 422
Db 291 GCGCTGCGGCTTGTGACGCTTCTCTTCTATGCAATCTACATGGCAGCATCTCTGTTCTCT 350
Qy 423 GACCTGCATCTCTGTGCACCACTGCTAGGTGTGTGGCCACCACTGTTGCTGGCTGCCCTA 482
```

```
Db 351 CACCTGATTAAGCTTCAGCGCTACCTGGGCATCTGCCACCCCTGGCTTCCTGGCAAA 410
Qy 483 CC---GGACCCGCGAGGATGCTGGCTGGGCGCCAGCACACCTGGGCGCTGGTGGTCT 539
Db 411 GCGTGGAGGTGCGCGTGTCTGCTGGGTAGTGTGTGGAGTCTGTGTGGTGGTGTGACAGC 470
Qy 540 CCAGTGTGCGCCACACTGGCCTTCTCCACACGAGCTACATCAATGGCCAGATGATCTG 599
Db 471 CCAGTGTGCGCCACGCGAGTCTTTGTGCCACAGGATCCAGCGCAACCCGACTGTGTG 530
Qy 600 GTATGACATGACCAAGCAAGAAATTTGATPGGCTTTTGGCTTACGGCATAGTCTGAC 659
Db 531 CTACGACCTGAGCCACCCATCTGTCTACTCGCTACCTGCGCTATGATGGCCCTCAC 590
Qy 660 ATTGTCTGGCT--TTCCTTCCCTCTTGGTCAATTTTGGTGTGCTATTCACTGATGCTCAG 717
Db 591 GGTCAATCGGCTTCTTGTGCTGCGCTTTCATAGCCCTTACTGGCTTGTATTGTCCCATGGCCCG 650
Qy 718 GAGCTGTATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGTC 777
Db 651 CCGCTGTGTGCCAGGATGGCCAGCAGGCTCTGTGSCCCCAAGAGCGCGCAGCAAGGC 710
Qy 778 CATCCGACCATCTCTACTGTGTGGCTCTTCACTCTGTGTGCTTGTGGCTTCCATAT 837
Db 711 GGTCTGTATGCTGTGTGTGGTGGCAGCTGTCTTTGCCATCAGCTTCTTGCCTTTCCACAT 770
Qy 838 CACTGCTCTTCTTACCTCACCATCTGCTTTTCTGCTTCTCAGGACTGCCAGCTCTTGAT 897
Db 771 CACCAAGACAGCTTACTTGGCTGTGCGCTCCAGCCCGGTGCTCTTGGCCCTGTGCTGGA 830
Qy 898 GGCAGCCAGTGTGGCTTACAAGATATGAGGCGCTTGTGTGAGTGTGAGCAGCTGCTCAA 957
Db 831 GACCTTCGCTGCTGCTGTACAAAGGCACTCGGCGCTTCGCCAGTGTCAACAGTGTCTGGA 890
Qy 958 CCCAGTCTCTGACT 971
Db 891 CCCCATCTCTTCT 904
```

Search completed: November 10, 2004, 11:51:16
Job time : 116 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2004, 08:39:02 ; Search time 3904 Seconds
(without alignments)
10099.331 Million cell updates/sec

Title: US-10-763-972-1
Perfect score: 1082
Sequence: 1 atgctgtccatttgcctcc.....ccagggttgacagatctgg 1082

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_est4.*
- 5: gb_est5.*
- 6: gb_est6.*
- 7: gb_est7.*
- 8: gb_est8.*
- 9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	200.6	18.5	2542	3 AK017378	Mus muscu
2	200.6	18.5	3001	3 AK005013	Mus muscu
3	177.6	16.4	693	6 CD217676	pgrln.pk0
4	163.2	15.1	987	9 AY400797	Homo sapi
5	163.2	15.1	1425	3 CR624871	full-leng
6	163.2	15.1	1515	3 CR612681	full-leng
7	163.2	15.1	1532	3 CR618945	full-leng
8	163.2	15.1	1564	3 CF626266	full-leng
9	163.2	15.1	1589	3 CR626754	full-leng
10	163.2	15.1	1592	3 CR605588	full-leng
11	163.2	15.1	1594	3 CR605834	full-leng
12	163.2	15.1	1638	3 CR617080	full-leng
13	163.2	15.1	1651	3 CR597500	full-leng
14	163.2	15.1	1686	3 CR626706	full-leng
15	163.2	15.1	1690	3 CR598765	full-leng
16	161.8	15.0	925	5 BX704227	BX704227
17	158.8	14.7	4439	3 AK076364	Mus muscu
18	155.6	14.4	987	9 AY400798	Pan trogl
19	148.8	13.8	987	9 AY400799	Mus muscu
20	143.6	13.3	639	6 CA352828	624212 NC
21	141.2	13.0	641	5 BQ396255	NISC tg19
22	141.2	13.0	1090	5 BX397648	BX397648
23	140.2	13.0	627	1 AL675845	AL675845
24	140.2	13.0	1168	1 AL561864	AL561864

C	25	140	12.9	1023	5	BX332649	BX332649
	26	134.8	12.5	1815	7	CN210812	RJA041C08
	27	133.6	12.3	1013	3	CR614658	full-leng
C	28	131	12.1	1057	5	BX381744	BX381744
	29	128.4	11.9	1090	3	CR598708	full-leng
	30	128.2	11.8	899	5	BQ591110	AGENCOURT
C	31	127.2	11.8	1095	5	BX358251	BX358251
	32	126	11.6	1035	5	BX379990	BX379990
	33	125.8	11.6	1042	5	BX397649	BX397649
	34	124.4	11.5	828	6	CD110272	CD110272
	35	123.2	11.4	623	5	BQ038875	BQ038875
	36	123.2	11.4	1065	5	BX337204	BX337204
	37	122.8	11.3	993	5	BX394202	BX394202
	38	122.4	11.3	955	5	BX354044	BX354044
	39	122	11.2	970	5	BX385626	BX385626
	40	121.6	11.2	851	4	B1833118	B1833118
	41	120.6	11.1	1076	5	BX375199	BX375199
	42	120.2	11.1	1063	4	BN918491	AGENCOURT
C	43	119.8	11.1	408	4	B1401676	B1401676
	44	119.8	11.1	747	5	BX922819	BX922819
C	45	119.6	11.1	1101	9	CNS05325	AL318925

ALIGNMENTS

RESULT 1
AK017378 2542 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15 product:PURINERGIC RECEPTOR P2Y, G-PROTEIN COUPLED 2, full insert sequence.
DEFINITION AK017378.1 GI:12856588
ACCESSION AK017378
VERSION HTC; CAP trapper.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
Carninci, P. and Hayashizaki, Y.
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
PUBMED 11076861
REFERENCE 5
The FANTOM Consortium and the RIKEN Genome Exploration Research

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9279253 PUBLISHED
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3001)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Atai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imokani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

On Dec 10, 2002 this sequence version replaced gi:12836638.
please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGCGCGCGACACGAGTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using trichloroethane thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'

GAGAGAGAGAGATCCAAAGAGCTCAATTAATTAATAACCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

FEATURES
source Location/Qualifiers
1. .3001
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:1300015C04"
/db_xref="taxon:10090"
/clones="1300015C04"
/sex="male"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
477..1598
/note="unnamed protein product; PURINERGIC RECEPTOR P2Y, G-PROTEIN COUPLED 2 (SPTR|Q9CP24, evidence: FASTY, 100%ID, 100%length, match=1119)
putative"
/codon_start=1
/protein_id="BAB23746.1"
/db_xref="GI:12836639"
/translation="MAADLEPNSTINGWEGDELGVKCRNEDFKVLLPVSYGVVC
VLGLANVALYIFLCLKLTWASTTVMFLAVSDSLVYASLPVLYVYAGDHWPF
TVLCQVRFYTNLYCSILFTICISVHRCGLVLRPLSLWGRVARRVAAVWVL
VIACQPVLYFTVSVRGTRITCTDTSARELFHSFVAYSSVNLGILFAPVPSVLYCY
VIMARLLKPAYTGGTGLPRAKRSVRIALVAFVLCFLPFLHTRILYISFSLDL
SCHTLNAINMAYKITRPLASNSCLDPVLYELAGORLVRFDARDAKPEPTPTSPQARR
KLGLHRPNRTVRKOLSVSSDDSRRTSTPAGSETKDIRL"

2982..2987
/note="putative"
3001
/note="putative"

polyA_signal
polyA_site

ORIGIN

Query Match 18.5%; Score 200.6; DB 3; Length 3001;
Best Local Similarity 54.5%; Pred. No. 3.3e-43;
Matches 469; Conservative 0; Mismatches 384; Indels 8; Gaps 3;

QY 127 TCCTGAGTTCACAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186
Db 549 TGTCTGTTTCAACGAGGACTTCAAGTACGTGCTGTTTGGCCGTCTCTATGGCGTGTGTC 608

QY 187 ATCTAGAGGCTGCCACTAAATGGCACTCTCTGTGGCACTCTCTGGGCACTCTCTGGGCAACACCAAGCGC 246
Db 609 GTGCTCGGCTTGTGCTGAACTCTGTGCTCTCTATATCTTCTATGCGCTCTAAACC 568

QY 247 TGGAGCTGTGCCACCACTATCTGCTGAACCTGTATGCTGGCGGCACTCTCTTTATGTGCTA 306
Db 669 TGGACGCTTCCACCACTACATGTTTCACTGGCAGTTTGGACTCTCTCTACGAGCG 728

QY 307 T---TGCCCTTCCCTCATCATCACTACTACTAGATGACAGTGGCCCTTCGGGAGCTG 363
Db 729 TCCCTGCCCTTGTGTTTATTACTACGCCGGGTGACCACTGGCCATTTAGCAGCGTG 788

QY 364 CTCTGCAAGCTGGTGCACTTCTCTTCTATATCAACCTTTTACGGCAGCATCTCTGCTGTCG 423
Db 789 CTCTGCAAGCTGGTGCGTTTCTCTTCTACACCACTCTACTGAGCATCTCTCTTCTCTC 848

QY 424 ACTTGCACTCTGTGCAACCACTTCTAGGTGTGTGCCACCACTGTGTGCTGCTGCCCTAC 483
Db 849 ACCTGATCAGCGTGCACCGGTGCTGAGTCTCTGGCCCTCTGCACCTCCCTGGGTGG 908

QY 484 CGGACCGGAGCATGCTGCTGGTGGCAACCACTGCTGGTGGTCTGCTCTCCAG 543
Db 909 GGGCGGGCGGTGTATGCCCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 968

QY 544 CTGCTGCCACACTGGCCCTTCTCCACACGCACTACATCAATCAATGGCCAGATGATCTGGTAT 603
Db 969 CAGGCACCGTGTCTCTACTTGTGTCAACCACTGCGGGGAAACCGGATCACTTGGCCAT 1028

QY 604 GACATGACCCAGCCAGAGATTTTGATCGCTTTTTCCTACGGCATAGTCTCTG--ACAT 661

```

|||||
1029 GACACCTGGCCGAGAGCTCTTTAGCATTTTGGCTTACAGCTCGTATCGGT 1088
QY 662 TGCTCTGGCTTTCTTTCCCTCTGTGTCTATTTGGTGTCTATTCACATGATGTCAGGAGC 721
Db 1089 CTGCTTTTGTGTGCTTTCCGTAATCTCTGTCTGTACGTGCTTATGGCCAGCGG 1148
QY 722 CTGATCAAGCCAGAGAGAACTCATAGAGAGAGCAACAGCCCGAGCCAGGTCCATC 781
Db 1149 CTGCTCAAAACCGGCTTATGGGACCAAGAGAGGTCTGCTCGGGGCAAGCGCAAGTCTGTG 1208
QY 782 CGGACCATCTACTGCTGTGGGCTCTTACCTCTGTTTGTGGCCCTTCCATATCACT 841
Db 1209 CGACCATGCTGTGTACTGGCGCTCTTCCCTCTGCTTCTGCTTCCCTTCCACTCAG 1268
QY 842 CGCTCTCTTACTACCATGCTGTTTCTGCTTCTCAGGACTGCCAGCTTCTGATGGCA 901
Db 1269 CGACCTCTACTACTCTCTCCGAT---CACTTGACCTCAGCTGCCACACCTCAAGGCC 1325
QY 902 GCGAGTGGGCTACAAGATATGGAGGCTCTGCTGTGAGTGTGACAGCTGCTCAACCCA 961
Db 1326 ATCAACATGGATATAGATCACCGGCGCTGGCCAGCGCAACAGTTGCTTGGCCCG 1385
QY 962 GTCCTGTACTTTCTTCAAG 982
Db 1386 GTACTCTACTTCTTGGCAGGG 1406

RESULT 3
CD217676
LOCUS
DEFINITION
pgrln.pk004.g6 Normalized chicken reproductive tract cDNA library
(pgrln) Gallus gallus cDNA clone pgrln.pk004.g6 5' similar to
sp|Q98907|P2Y3_CHICK_P2Y purinoceptor 3 (P2Y3) (Nucleoside
diphosphate receptor) >emb|CAA65930.1| (X98283) G protein-coupled
P2 receptor (Gallus gallus), mRNA sequence.
CD217676
ACCESSION
CD217676.1 GI:30957649
VERSION
EST.
KEYWORDS
Gallus gallus (chicken)
SOURCE
Gallus gallus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 693)
Cognburn, L.A. and Nys, Y.
Chicken ESTs from reproductive tract
Unpublished (2002)
Contact: Larry A. Cognburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cognburn@udel.edu, www.chickest.udel.edu.

FEATURES
source
1..693
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Commercial broiler and layer"
/db_xref="taxon:9031"
/clone="pgrln.pk004.g6"
/sex="Male and Female"
/tissue_type="Testis, ovary and oviduct"
/dev_stage="Various stages; embryonic, post-hatch, immature
and sexually-mature"
/lab_host="E. Coli BMDH10B"
/clone_lib="Normalized chicken reproductive tract cDNA
library (pgrln)"
/note="vector: pCMVSPORT6; Library made from three total
RNA pools from each tissue (testis 25%, ovary 25%, and
oviduct 50% of final RNA pool) Single pass sequencing from
5'-end"


```

```

Query Match 16.4%; Score 177.6; DB 6; Length 693;
Best Local Similarity 59.2%; Pred. No. 4.2e-37;
Matches 359; Conservative 0; Mismatches 239; Indels 8; Gaps 3;

QY 127 TGGCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186
Db 67 TGCACCTTCATGAGAAATTCAGGAGGTCCTGCTGCCCTGCTACTCAGTGTGTTC 126
QY 187 ATCTAGGGCTCCCACTAAATGGCACTGTCTTTGTGCGCACTCCTGGGGGCCAACAAGGCG 246
Db 127 CTACTGGGCTCGCACTCAATGCCGTGTCTATTGGCGAGATCTGCTGGCCGCAAGCGG 186
QY 247 TGGAGTGTGGCACACACATATCTGTGAACCTGTGTTGGCGGACCTGCTTTATGTGCTA 306
Db 187 TTGAOCGCGACCAACCATCTACATGCTGAACCTGGCCATGGCGGACCTGCTTTATGCTGC 246
QY 307 T---TGCCCTTCTCATCATCACTACTCACTAGATGACAGGTGGCCCTTCGGGGAGCTG 363
Db 247 TCCCTCCCTCTCTCATCTACACTACACCCAGAGGATTACTGGCCCTTTGGGACTTC 306
QY 364 CTCTGCAAGCTGGTGACCTTCTGTTCTATATCAACCTTTTACGGCAGCATCTCTGCTGTG 423
Db 307 ACCTGCAAAATTCGTCGGTTCAGTTTCTACACCAACCTGCGGCGAGCATCTCTTCTCTC 366
QY 424 ACCTGCATCTGTGTGACACAGTTCCTAGGTGTGTGCCACCCACTGTGTTGCTGCCCTAC 483
Db 367 ACCTGCATCAGCTCCAGGCTACATGGGATCTGCCACCCCTTGGCCCTGTGGCAGAA 426
QY 484 CGGACCCG---CAGGCATGCTGCTGGCGGACACCAACCACTGGGCCCTCTGCTCTCTC 540
Db 427 AAGAAGGAAAGAAAGCTCAGCTGGCTGTGTGCTGCTGCTGTTTTCATGTCATCGCC 486
QY 541 CAGCTGTGCTGCCACACTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 600
Db 487 CAGTGCCTGCCACACTTTGTTCTTCCCTCCACCGGACGCGAGAGATTCGCACTGTCTGC 546
QY 601 TATGACATGACGACGACAGAGAAATTTGATCGGCTTTTGTGCTACGGCATAGTTCTTGACA 660
Db 547 TATGACCTGAGCCCCCGGACCGCTCCACATCTTCCCTATGGCATCAGCTTGACC 606
QY 661 TTGCTGTGGCTTCT--TTCCCTCTCTGTGCTATTTGGTGTGCTATTTACTGATGTGAGG 718
Db 607 ATCACCGGCTTCTGCTGCTCCCTTCGACCCATCCTGGCCCTGCTACTGCGATGCGCCGC 666
QY 719 AGCCTG 724
Db 667 ATCTCTG 672

RESULT 4
AY400797
LOCUS
DEFINITION
Homo sapiens P2RY6 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY400797
VERSION
AY400797.1 GI:39756786
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 987)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 987)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A.,

```

Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferrira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment:
Location/Qualifiers
1..987
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..987
/gene="P2RY6"
/locus_tag="HOM0677"

Query Match 15.1%; Score 163.2; DB 9; Length 987;
Best Local Similarity 51.9%; Pred.No. 4.1e-33;
Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;

Qy 126 CTGCCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTT 185
Db |||||
Qy 51 CTGTGTCTACCGCAGAACTCTCAGCAACTGCTGCTCCACCTGTGTATTTCGGCGGTGCT 110
Db |||||
Qy 186 TATCCTAGGGTGCACCTAAATGGCACTGCTTGTGGCACTCTCTGGGGCCAAACCAAGCG 245
Db |||||
Qy 111 GGCGGCTGGCTGGCTGGAACATCTGTGTGTCATTACCCAGATCTGCAGCTCCGCGGGC 170
Db |||||
Qy 246 CTGGAGCTGTGCCACCACTATCTGTGTGAACCTGTAGTGGCCGCACTGCTTTATGTGCT 305
Db |||||
Qy 171 CCTGACCGCAGCGCCGCTGACACCTTAACTTGTCTGTGCTGACCTGTATATGCTG 230
Db |||||
Qy 306 AT---TGCCCTCTCTCATCATCTACTACTACTAGATGACAGTGGCCCTTCGGGAGCT 362
Db |||||
Qy 231 CTCCCTGCCCTGCTCTATCAAACTATGCCAAGGTGATCAGTGGCCCTTTGGCGACTT 290
Db |||||
Qy 363 GCTCTGAAGTGTGCATCTCTCTTCTATATCAACCTTTACGCGAGCATCTCTGTGCT 422
Db |||||
Qy 291 CGCTGCGCGTGTGCTGCTTCTTCTATGCAACCTTCACGCGAGCATCTCTTCTCT 350
Db |||||
Qy 423 GACTGTATCTCTGTGACCAAGTCTCTAGTGTGTGTCACCCACTGTGTCGTGCGCTA 482
Db |||||
Qy 351 CACCTGTATCAGTTCAGCGCTTACCTTGGGCACTGCGCCGCTGGCCCTTGGCACA 410
Db |||||
Qy 483 CC---GGACCGCAGGCGATGCTGGCTGGGCAACAGCACCACTGGCCCTGTGGTGTCT 539
Db |||||
Qy 411 ACCTGGGGGCGCGGGCTGCTGGCTAGTGTGTGTAGCGTGTGGCTGGCGGTGACAAC 470
Db |||||
Qy 540 CCAGCTGTGCCCCACACTGGCCCTTCTCCACAGCACTACATCAATGGCCAGATGATCTG 599
Db |||||
Qy 471 CCAGTGTGCGCCACAGCAGCATCTTGGCTGCGCCACAGGCATCCAGCGTAACCGCACTGCTG 530
Db |||||
Qy 600 GTATGATGATGACGAGGAGATTTTGTATGGCTTTTGGCTTACGGCATAGTCTGAC 659
Db |||||
Qy 531 CTATGACCTCAGCCCGGCTGCGCTGGCCACCACTATATGCCCTATGGCATGGCTCTCAC 590
Db |||||
Qy 660 ATTGTCTGGCTTCTCT---TTCCCTCTCTGGTCAATTTTGGTGTGCTATTACTGATGTGTAG 717
Db |||||
Qy 591 TGTATCGGCTTCTGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650
Db |||||
Qy 718 GAGCTGTATGACGAGGAGAGACTCATGAGAGAGGAGCAACACAGCCGAGCGCAGGTC 777
Db |||||
Qy 651 CCGCTGTGCGCGCAGGATGCGCGGAGAGCCCTGTGGCCAGAGGCGGCTGGCAAGGC 710
Db |||||
Qy 778 CATCGGAGCATCTACTGTGTGTGGCTCTTTCACCTCTGTTTGTGGCCCTTCCATAT 837
Db |||||
Qy 711 GGCCCGCATGGCGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 770
Db |||||
Qy 838 CACTGTCTCTTCTACCTCAACAATGCTGCTTCTGCTTCTTCTCAGGAGCTGCCAGCTCTTGTAT 897
Db |||||
Qy 771 CACCAAGACAGCCTACCTGGCAGTGGCTCGAGCCGCGGCTGCCCTGCTGCTGCTGCTGCTG 830
Db |||||

Qy 898 GGCACCCAGTGTGGCTTACAGATATGAGGCTCTCTGCTGAGTGTGACGAGCTGCTCAA 957
Db |||||
Qy 831 GGCCTTTTTCAGCGGCTTACAAAGGACCGCGGCTTTCCAGTGCACACAGGCTCTGGA 890
Db |||||
Qy 958 CCCAGTCTCTGTACT 971
Db |||||
Qy 891 CCCCATCTCTTCT 904
Db |||||

RESULT 5
LOCUS CR624871 1425 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0D1035YK19 of Placenta Cot 25-normalized
of Homo sapiens (human).

ACCESSION CR624871
VERSION CR624871.1 GI:50505678
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1425)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished

REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1425)
Genoscope.

REFERENCE Direct Submission
AUTHORS Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :
TITLE Bp 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
JOURNAL - Web : www.genoscope.cns.fr

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES Location/Qualifiers
source 1..1425
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1035YK19"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 15.1%; Score 163.2; DB 3; Length 1425;
Best Local Similarity 51.9%; Pred.No. 4.5e-33;
Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;

Qy 126 CTGCCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTT 185
Db |||||

Qy 224 CTGTGTCTACCGCAGAACTTCAAGCACTGCTGTGCCACTGTGTATTCGCGGTGCT 283
Db |||||

Qy 186 TATCCTAGGGTGCACCTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCG 245
Db |||||

Qy 284 GGCGGCTGGCTGCGCTGCAACATCTGTGTCTATACCCAGATCTGCACGTCGCCGCGGCG 343
Db |||||

Qy 246 CTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCGGCTGCTTTATGTGCT 305
Db |||||

Qy 344 CTGACCCCGCAGCGCGGTGTACCCCTAACTTCTGCTGGCTGACCTGTATATGCTG 403
Db |||||

Qy 306 AT---TGCCCTTCTCTCATCATCACCTACTCTACTAGATGACAGTGGCCCTTCGGGAGCT 362
Db |||||

Qy 404 CTCCTGCCCCCTGCTCATCTACAATATGCCCAAGTGTATCATCTTTTGGGACTT 463
Db |||||

Qy 363 GCTCTGCAAGCTGTGTGCACTTCTGTTCTATATCAACCTTTACGCGAGCATCTGCTGCT 422
Db |||||

Qy 464 CGCTCGCCGCTGGTCCGCTTCTCTTCTATGCCAACTTGCACGCGAGCATCTCTCTCT 523
Db |||||

```
QY 423 GACCTGATCTCTGTGACAGTTCCTAGTGTGTGTGACACCACTGTGTGTGCTGACCTA 482
Db      |||
QY 524 CACCTGATCATGCTTCCAGCGCTACCTGGGATCTGTGACACCGCTGGCCCTGGCACAA 583
Db      |||
QY 483 CC---GGACCCGAGGATGCTGTGCTGGGACACGACACCACTGGGCCCTGGTGGTCCCT 539
Db      |||
QY 584 AGTGGGGGCGCGGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 643
QY 540 CCAGCTGTGCCCCACATGCGCTTCTCCACACGAGTACATCAATGGCCAGATGATCTG 599
Db      |||
QY 644 CCAGTGGCTGCCCCACAGCATCTTGGCTGCCACAGGATCCAGCGTAAACCGCACTGTCTG 703
QY 600 GTATGACATGACCGACAGGAGAAATTTGATCGGCTTTTTCCTACCGCATAGTTCGAC 659
Db      |||
QY 704 CTATGACCTGACCGCGCTGCGCTGGCCACCACTATATGCGCTATGGCATGGCTCTCAC 763
QY 660 ATTGTCTGGCTTTCT--TTCCCTCTCTGGTCAATTTGGTGTGTATTTCACTGATGTTCAG 717
Db      |||
QY 764 TGTATCGGCTTCTGTCTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 823
QY 718 GAGCTGTATCAAGCAGAGGAGACCTCATGAGCAGAGGACACACGCGCCGAGGCTC 777
Db      |||
QY 824 CCGCTGTGCGCCGAGGATGCGCCGAGAGCCCTGTGCGCCAGAGCGGGCTGGCAAGGC 883
QY 778 CATCGGACCATCTACTGTGTGTGTGCGCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 837
Db      |||
QY 884 GGCCGCGATGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 943
QY 838 CACTGCTCTCTTACCTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
Db      |||
QY 944 CACCAAGACAGCTACCTGGCAGTGGCTGCGCTGCGCGCGGGCTGCCCTGCACTGTATTGGA 1003
QY 898 GCGAGCCAGTGTGGCTACAGATATGAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 957
Db      |||
QY 1004 GGCCTTTCAGCGGCTTCAAGGACGCGCGCGCTTTCAGAGTCCCAAGCGGCTGTGGA 1063
QY 958 CCCAGTCTCTGACT 971
Db      |||
QY 1064 CCCATCTCTTCT 1077
Db      |||

RESULT 6
CR612681
LOCUS      1515 bp      mRNA      linear      HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0D1044YA15 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR612681
VERSION CR612681.1 GI:50493488
KEYWORDS HTC; CNSLT_CDNA
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1515)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES Location/Qualifiers
```

```
source 1..1515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1044YA15"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 15.1%; Score 163.2; DB 3; Length 1515;
Best Local Similarity 51.9%; Pred. No. 4.6e-33;
Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;

QY 126 CTGCGAGTTCTCAGAGAAAGTACAAGCAAGTCTACCTCTCCCTGGGCTACAGTATCATCTT 185
Db      |||
QY 312 CTGTGTCTACCCGAGAACTTCAAGCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 371
Db      |||
QY 186 TATCTAGGGCTGCCACTAAATGGCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 245
Db      |||
QY 372 GCGCGCTGGCTGCGCTGAACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 431
QY 246 CTGAGCTGTGCCACACCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 305
Db      |||
QY 432 CTGACCCGCGACGGCGGTGTACCCCTAAACCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 491
QY 306 AT---TGCCCTTCTCTCATCATCACCTACTACTAGATGACAGGTGGCCCTTCGGGAGCT 362
Db      |||
QY 492 CTCCCTGCCCTGCTCATCTACCACTATGCCAAGTGTACATGCGCCCTTTGGCGACTT 551
QY 363 GCTCTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 422
Db      |||
QY 552 GCGCTGCGCGCTGGTCCGCTTCTCTTATGCCAACCTGTGACCGGACGATCCCTTCTCT 611
QY 423 GACCTGTGATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 482
Db      |||
QY 612 CACCTGTGATCAGTTCAGGCTACTTGGGATCTGTGCACTGTGTGTGTGTGTGTGTGTGTGTGT 671
QY 483 CC---GGACCCGAGGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 539
Db      |||
QY 672 AGTGTGGGCGCGCGGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 731
QY 540 CCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 599
Db      |||
QY 732 CCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 791
QY 600 GTATGACATGACACGCAAGAGAAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
Db      |||
QY 792 CTATGACCTCAGCGCGCTGCGCTGGCCACCACTATATGCGCTATGCGCTATGGCATGGCTCTCAC 851
QY 660 ATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 717
Db      |||
QY 852 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 911
QY 718 GAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 777
Db      |||
QY 912 CCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 971
QY 778 CATCGGACCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 837
Db      |||
QY 972 GCGCGCGATGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1031
QY 838 CACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 897
Db      |||
QY 1032 CACCAAGACAGCTTACCTGGGAGTGGCTTGCAGCGCGGGCGTCCCTGCACTGTATTGGA 1091
QY 898 GCGACGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 957
Db      |||
QY 1092 GCGCTTGTGAGCGGGCTCAAGAGGACGCGCGCGCTTTCAGAGTGCACACAGCGGTGTGTGGA 1151
QY 958 CCCAGTCTCTGACT 971
Db      |||
QY 1152 CCCATCTCTTCT 1165
Db      |||
```

```

RESULT 7
CR618945
LOCUS
DEFINITION
    1532 bp mRNA linear HTC 21-JUL-2004
    full-length cDNA clone CS0D1042YA16 of Placenta Cot 25-normalized
    of Homo sapiens (human).
ACCESSION
CR618945
VERSION
CR618945.1 GI:50499752
KEYWORDS
HTC; CNSLT_cDNA.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1532)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1532)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1042YA16"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 15.1%; Score 163.2; DB 3; Length 1532;
Best Local Similarity 51.9%; Pred. No. 4.6e-33;
Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;
QY 126 CTGCAGTCTTCAGAGAGTACAGCAAGTCTACTCTCCCTGGCCCTACAGTATCATCTT 185
DB 312 CTGTCTACCGCGAGAACTTCAGCAACTGCTCTGCCACCTGTGTATTCCGGGGTCT 371
QY 186 TATCCTAGGGTGCCACTAATATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCG 245
DB 372 GCGCGTGGCTGGCGCTGAACATCTGTCTATACCCAGATCTGCACGCTCCGCGCGGC 431
QY 246 CTGGAGCTGTGCCACCACTATCTGGTAACCTGATGGTGGCGACCTGCTTTATGTGCT 305
DB 432 CTTGACCGGACGCGCGGTGTACACCTTAACCTTGTCTGTGGTACCTGTATATGCTG 491
QY 306 AT---TCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTGGGGAGCT 362
DB 492 CTCCTGCGCCCTGTCTATCACTAATACTATGCCAAAGTGTACCTGGCCCTTGGCGACT 551
QY 363 GCTCTGAGCTGTGCTACCTCTCTCTATATCAACCTTTACGGCAGCATCTCTGCTGCT 422
DB 552 CGCTTGGCGCTGGTGGCTTCTCTCTATATGCAACCTTGACGGCAGCATCTCTCTCT 611
QY 423 GACCTGATCTCTGTGCCACCACTTCTAGGTGTGTGGCACCCTAGTGTTCGTGGCCCTA 482
DB 612 CACCTGATCAGCTTCCAGCGCTACCTGGGATCTGCCACCGCTGGCCCGCTGGCAAA 671
QY 483 CC---GGACCGGAGGCGATGCTGGCTGGGACACAGACACCTCTGGCCCTGTGTGCTCT 539
DB 672 ACCTGGGGGCGCGCGGCTGCTGCTAGTGTGTGTAGCCGTGTGGCTGGCGGTGACAC 731

```

```

QY 540 CCAGCTGCTGCCACACATGGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTG 599
DB 732 COAGTGCTTGGCCACAGCCATCTTCGTGCGCACAGGCAATCCAGCGTAACCGCACTGTCTG 791
QY 600 GTATGACATGACCGACCGCAAGAGAAATTTGATCGGCTTTTGGCTTACGGCATAGTCTGAC 659
DB 792 CTATGACCTCAGCGCGCTGCGCTGGCCACCACTATATGCGCTATGGCATGGCTCTCAC 851
QY 660 ATTGTCTGGCTTCT--TTCCCTCCTTGGTGTATTTGGTGTGCTTATTCACATGATGGTCTG 717
DB 852 TGTATCGGCTTCTTGTGCTGCTTGTGCGCTTGTGCGCTTGTGCTTCTCTCTGGCGCTG 911
QY 718 GAGCCTGATCAAGCCAGAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCGAGTCT 777
DB 912 CGCGCTGTGCGCGCAGGATGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 971
QY 778 CATCCGAGCACTTCTACTGCTGTGTGGCTCTTACCTCTCTGTTTGGCCCTTCCATAT 837
DB 972 GCGCGCATGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1031
QY 838 CACTGCTCTCTTCTACTCATCATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 897
DB 1032 CACCAAGACAGCTTACTTGGCAGTGGCTCGACGCGCGGCGTCCCTGCTGCTGATTGGA 1091
QY 898 GGCAGCCAGTGTGGCTTACAGATATGAGGCTCTGTGAGTGTGAGCAGCTGCTCAA 957
DB 1092 GGCCTTTCAGCGGCTTACAAAGGACGCGCGGCTTTCGCACTGTCGCAACAGCGTCTGGA 1151
QY 958 CCCAGTCTCTGACT 971
DB 1152 CCCATCTCTTCT 1165
RESULT 8
CR626266
LOCUS
DEFINITION
    1564 bp mRNA linear HTC 21-JUL-2004
    full-length cDNA clone CS0D1038YC15 of Placenta Cot 25-normalized
    of Homo sapiens (human).
ACCESSION
CR626266
VERSION
CR626266.1 GI:50507073
KEYWORDS
HTC; CNSLT_cDNA.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1564)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1564)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1564
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1038YC15"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN

```

```

Query Match      15.1%; Score 163.2; DB 3; Length 1564;
Best Local Similarity 51.9%; Pred. No. 4.7e-33;
Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;

QY 126 CTGCCAGTTCTCAGAGAGTCAAGCAAGTCTACTCTCCCTGGCCTACAGTATCATCTT 185
DB |||
DB 343 CTGTGTCTACCGGAGAACTTCAAGCAACTGCTGCCACTGTGTATTGGGGGTGCT 402
QY 186 TATCTAGGGTGCACATAAATAGGCACTGTCTTGGCACTCTCTGGGGCCAAACCAAGCG 245
DB |||
DB 403 GCGGCTGCGCTGCGCTGGAACATCTGTGTCTATTACCCAGATCTGCACCTCCCGCGGC 462
QY 246 CTGGAGCTGTGCCACCACTATCTGTGAACCTGATGTGGCCGACCTGCTTTATGTGCT 305
DB |||
DB 463 CTGACCCGCGAGCGGTGTACACCTTAACCTTGTCTGGCTGACCTGCTATATGCTG 522
QY 306 AT---TGCCCTTCTCATCATCACTCACTAGTACAGAGTGGGCCCTTGGGGAGCT 362
DB |||
DB 523 CTCCCTGCCCCGTGTCTCATCACTATGCCAAAGGTGATCACTGGGCCCTTTGGCGACTT 582
QY 363 GCTCTGCAAGTGTGTGCACTTCTCTTCTATATCAACCTTTACGCGAGCATCTGCTGCT 422
DB |||
DB 583 GCGCTGCGCTGTGTGCTCTCTTCTATGCCAATCTGCACGCGAGCATCTCTTCTCT 642
QY 423 GACCTGCATCTCTGTGCACTAGTCTCTAGTGTGTGCAACCACTGTGTGCTGCCCCA 482
DB |||
DB 643 CACCTGCATCACTGTCCAGCGTACTTGGGCACTTGCACACCGGCTGGCCCTGGCACA 702
QY 483 CC---GGACCCGCGAGCATGCTGCTGGGCAACAGCAGCACCACCTGGGCCCTGTGCTCT 539
DB |||
DB 703 ACGTGGGGCGCGCGGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 762
QY 540 CCAGCTGTGCTGCCACACCTGCTCTCCCAACGCACTATCAATCAATGGCCAGATGATCTG 599
DB |||
DB 763 CCAGTGCCTGCCCCAGCCATCTTGGCTGCCAGAGCATCCAGGTAACCGCACTGTCTG 822
QY 600 GTATGACATGACGAGGAGAACTTGTGATCGGCTTTTGGCTTACGCGCATAGTTCTGAC 659
DB |||
DB 823 CTATGACCTCAGCCGCGCTGCGCTGGCCACCACTATATGCCCCATGCGCATGGCTCTCAC 882
QY 660 ATTGTCTGCTTCT---TTCCCTCTCTGCTATTTGGTGTGCTTATTCATGATGCTGTCAG 717
DB |||
DB 883 TGTATCGGCTTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 942
QY 718 GAGCTGTATCAAGCAGGAGGAACTCATGAGGAGCAAGCAACACAGCCGAGCCAGGTC 777
DB |||
DB 943 CCGCTGTGCGCCAGGATGCGCGGAGAGCTGTGCGCCAGAGCGGCTGCGCAAGGC 1002
QY 778 CATCGGACCATCTACTGCTGTGTGCTCTTCCACCTCTGTTTGTGCTTCCATAT 837
DB |||
DB 1003 GGCCGCGATGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1062
QY 838 CACTGCTCTTCTACCTCACTATCTGCTTTTCTGCTTTCTCAGGACTGCCAGCTCTTGAT 897
DB |||
DB 1063 CACCAAGACAGCTTACCTGGCAGTGTGCTGCAAGCGCGGTCTCCCTGCACTGTATTGGA 1122
QY 898 GCGAGCCAGTGTGGCCTACAGATATGGAGCCCTCTGCTGTGTGTGTGTGTGTGTGTGTGT 957
DB |||
DB 1123 GGCCTTTGACGCGGCTTACAAAGGAGCAGCGCGCGCTTTTGGCAGTGCCAAAGCGTGTGGA 1182
QY 958 CCCAGTCTCTGTACT 971
DB |||
DB 1183 CCCATCTCTTCT 1196

```

```

RESULT 9
CR626754
LOCUS CR626754 1589 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSOD1024VE14 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR626754
VERSION CR626754.1 GI:50507616
KEYWORDS HTC; CNSUT_cDNA.

```

```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1589)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
REFERENCE 2 (bases 1 to 1589)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1024VE14"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 15.1%; Score 163.2; DB 3; Length 1589;
Best Local Similarity 51.9%; Pred. No. 4.7e-33;
Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;

QY 126 CTGCCAGTTCTCAGAGAGTCAAGCAAGTCTACTCTCCCTGGCCTACAGTATCATCTT 185
DB |||
DB 382 CTGTGTCTACCGGAGAACTTCAAGCAACTGCTGCCACTGTGTGTATTGGGGGTGCT 441
QY 186 TATCTAGGGTGCACATAAATAGGCACTGTCTTGGCACTCTCTGGGGCCAAACCAAGCG 245
DB |||
DB 442 GCGGCTGGCTGCGCTGCACTCTGTGTCTATACCAGATCTGCAGTCCCGCGGCG 501
QY 246 CTGGAGCTGTGCCACCACTATCTGTGTGAACCTGTATGTGGCGGACCTGTCTTTATGCT 305
DB |||
DB 502 CTTGACCCGCGACGCGGTGTACACCTTAACCTTGTCTGCTGCTGCTATATGCTGCT 561
QY 306 AT---TGCCCTTCTCATCATCACTACTAGTACAGAGTGGCCCTTGGGGAGCT 362
DB |||
DB 562 CTCCCTGCCCCGTGTCTCTCTTCTATGCCAACTGTGCCAGGATCATCTGCGCCCTTTGGGACTT 621
QY 363 GCTCTGCAAGTGTGTGCACTTCTCTTCTATATCAACCTTTACGCGAGCATCTCTGCT 422
DB |||
DB 622 GCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
QY 423 GACCTGCATCTCTGTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
DB |||
DB 682 CACCTGCATCACTGTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
QY 483 CC---GGACCCGCGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
DB |||
DB 742 AGTGGGGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
QY 540 CCAGCTGTGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
DB |||
DB 802 CCAGTGTGCTGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861
QY 600 GTATGACATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659
DB |||
DB 862 CTATGACCTCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 921
QY 660 ATTGTCTGCTTCT---TTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717

```


Db 595 CTCCTCGCCGCTGCTATCTAATACTATGCCAAGGTGATCACTGGCCCTTTGGCGACTT 654
 QY 363 GCTCTGAAGCTGTGTGCACTTCTCTGTTCTATATCAACCTTTTACGGCAGCATCTCTGCT 422
 Db 655 GCGCTGCGCGCTGTGTCGCTTCTCTTCTATATGCCAAGGTGATCACTGGCCCTTTCTCT 714
 QY 423 GACCTGATCTCTGTGCACTGCTTCTAGGTGTGTGCCAAGGTGATCACTGGCCCTTTCT 482
 Db 715 CACCTGATCTGCTTCTGCACTGCTTCTAGGTGTGTGCCAAGGTGATCACTGGCCCTTTCT 774
 QY 483 CC---GGACCCGCGAGGATGCTGCTGGGCAACAGCAGCAGCAGCAGCAGCAGCAGCAG 539
 Db 775 ACGTGGGGGCGCGGGCTGCTGGCTAGTGTGTAGCGGTGTGGCTGGCGGTGACAC 834
 QY 540 CCAGCTGTGCTGCCACACTGCTTCTTCCCAACGAGCTTACATCAATGCGCAGATGATCTG 599
 Db 835 CCAGTGTGCTGCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 894
 QY 600 GTATGACATGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 659
 Db 895 CTATGACCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 954
 QY 660 ATTGTCTGGCTTCT--TTCCCTCTCTTGGTCACTTGGTGTGATTTTCTGCTGATGATG 717
 Db 955 TGTATCGGCTTCTGCTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
 QY 718 GAGCTGATCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 777
 Db 1015 CCGCTGTGCGCGCAGGATGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1074
 QY 778 CATCGGACCATCTACTGTGTGTGCTTCTTCAACCTCTGTTTGTGCTTCTTCCATAT 837
 Db 1075 GCGCGCATGCGCTGT 1134
 QY 838 CACTGCTCTTCTACTTCACTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 897
 Db 1135 CACCAAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1194
 QY 898 GGCAGCAGCTGTGCTTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCAGCAGCAGCAG 957
 Db 1195 GGCCTTGTGAGCGGCTTACAAAGCAGCGGCGCTTTGCCAGTCCCAACAGCGTGTGGA 1254
 QY 958 CCCAGCTCTGTACT 971
 Db 1255 CCCATCTCTTCT 1268

RESULT 13
 CR597500
 LOCUS full-length cDNA clone CS0DC005YM02 of Neuroblastoma Cot
 DEFINITION 25-normalized of Homo sapiens (human).
 ACCESSION CR597500
 VERSION CR597500.1 GI:50478307
 KEYWORDS HTC; CNSLUT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1651)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 Contact: Feng Liang Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 2 (bases 1 to 1651)
 Genoscope.
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
 - Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen

FEATURES
 source Location/Qualifiers
 1..1651
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC005YM02"
 /issue_type="Neuroblastoma Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN
 Query Match 15.1%; Score 163.2; DB 3; Length 1651;
 Best Local Similarity 51.9%; Pred. No. 4.7e-33;
 Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;

QY 126 CTGCCAGTCTTCAGAGAGTACAGCAAGCTACCTCTCCCTGGCTACAGTATCATCTT 185
 Db 438 CTGTGTCTACCGGAGAACTTCAAGCACTGTGTGCCACCTGTATATTCGGCGGTGT 497
 QY 186 TATCTTAGGGCTGCCACTAAATGGCACTCTCTTGGCACTCTCTGGGGCAAAACAAGCG 245
 Db 498 GCGGGCTGSCCTGCCCTGAACATCTGTGTCTATTACCGAGATCTGCACGTCCCGCGGCG 557
 QY 246 CTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTTGGCGGACCTGTCTTATGTGCT 305
 Db 558 CTGTGACCCGACGGCGGTGTACACCTAAACCTTGTCTGGCTGACCTCTATATGCTG 617
 QY 306 AT---TGCCCTTCTTCATCATCATCACTACTACTAGATGACAGTGGCCCTTCGGGGAGCT 362
 Db 618 CTCCTTCCCTCTGTCTCACTACTATGATGCCAAGGTGATCACTGGCCCTTTGGCGACTT 677
 QY 363 GTCTGCAAGCTGTGTGCACTTCTGTCTTATATCAACCTTTACGGCAGCATCTCTGTCT 422
 Db 678 CGCCTGCGCCCTGTGTCGCTTCTCTTATGCCAACCTGACCGCAGCATCTCTTCTCT 737
 QY 423 GACCTGCATCTGTGCACTGTTCTAGGTGTGTGCCACCACTGTGTTCCTGCTGCTTA 482
 Db 738 CACCTGCATCACTTCCAGCGCTACTCTGGGCACTGTGCCACCGCGCTGGCCCCCTGGCAAA 797
 QY 483 CC---GGACCCGCGAGCATGCTGCTGGGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGTCT 539
 Db 798 ACGTGGGGGCGCGGGCTGCTGCTAGTGTGTAGCGGTGTGGCTGGCGGTGACAC 857
 QY 540 CCAGCTGTGCTGCCACACTGCGCTTCTCCCAACGAGCTATCATCAATGGCCAGATGATCTG 599
 Db 858 CCAGTGTGCTGCCACAGCCATCTTGGCTGCCAAGGATCCAGCGTAAACCGCACTGTCTG 917
 QY 600 GTATGACATGACAGCAGCAGAGAACTTCTGATCGGCTTTTGGCTTACGGCATAGTTCTGAC 659
 Db 918 CTATGACCTCAGCGCGCTGCCCTGCCACCACTATATAGCCCTATGGCATGGCTCTCAC 977
 QY 660 ATTGTCTGGCTTCT--TTCCCTCTCTGGTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 717
 Db 978 TGTCTATGCGCTTCTGTCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1037
 QY 718 GAGCTGTATCAAGCAGCAGAGGAGAACTCTATGAGGACAGCAGCAGCAGCAGCAGCAGGTC 777
 Db 1038 CCGCTGTGCCCGCAGGATGGCCCGCAGAGCTGTGGCCCGCAGGCGCGGTGGCAAGCG 1097
 QY 778 CATCGGACCATCTACTGGTGTGTGGCTCTTCCACCTCTGTCTTTGTGCTCTTCCATAT 837
 Db 1098 GCGCCGATGCGCGGTGTGTGGCTGTGCTCTTCCATCAGCTTCTGCTCTTCTTCTCATAT 1157
 QY 838 CACTGCTCTTCTACTCATCATCTGCTTCTTCTGCTTCTCTCAGGAGTGGCAGCTTCTGAT 897
 Db 1158 CACCAAGACAGCTTACCTGGCAGTGTGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAA 1217
 QY 898 GGCAGCAGCTGTGGCTTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAA 957
 Db 1218 GGCCTTGTGAGCGGCTTACAAAGCAGCGGCGCTTGCAGTGTGCCACAGCGTGTCTGGA 1277

```

QY 958 CCCAGTCTCTGACT 971
Db 1278 CCCATCTCTTCT 1291

RESULT 14
CR626706 1686 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DB006YE03 of Neuroblastoma Cot
DEFINITION 10-normalized of Homo sapiens (human).
ACCESSION CR626706
VERSION 1 (bases 1 to 1686)
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1686)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1686
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB006YE03"
/tissue_type="Neuroblastoma Cot 10-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 15.1%; Score 163.2; DB 3; Length 1686;
Best Local Similarity 51.9%; Pred. No. 4.7e-33;
Matches 443; Conservative 0; Mismatches 403; Indels 8; Gaps 3;

QY 126 CTGCCAGTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTT 185
Db 464 CTGTGTTACCCGAGAACTTCAAGCACTGCTGTCGCCACTGTGATTCGGGGTCT 523

QY 186 TATCTAGGCTGCCCAATAATGGCACTGTCTGTGGCACTCTCTGGGCACTCTCTGGGCGCAAAACGCG 245
Db 524 GCGGGCTGCGCTGCGCTGAACATCTGTGTCTATTACCCAGATCTGCACGTCCCGCGGCG 583

QY 246 CTGGAGCTGTGCACCACTATCTGGTGAACCTGATGTCGCGGCGGCTTTATGTGCT 305
Db 584 CTGTACCCGCGCGCGGTGTACACCTAAACCTGCTCTGGCTGACCTGTATATGCTGT 643

QY 306 AT---TGCCCTTCTTCATCATCACTACTCTACTAGATGACAGGTGGCCCTTCGGGGAGCT 362
Db 644 CTCCTGCCCCCTGCTCATCTCACTCAATGTCGCCAAGTATCACTGCGCTTTGGCGACTT 703

QY 363 GCTCTGAAGCTGGTGCACTTCTGTTTATATCAACCTTTACGCGACATCTGCTGCT 422
Db 704 CGCCTGCGCCCTGGTCCGCTTCTCTCTATGCGCAACCTGCACGCGCATCTCTTCTCT 763

QY 423 GACCTGCATCTCTGTGACACAGTCTCTAGGTGTGTGGCAACCCACTGTGTGCTGCGCCCTA 482
Db 764 GACCTGCATCTCTGACCGCTACTCTGGGCACTCTGCCACCGCTGGCCCCCTGGCAAA 823

483 CC---GGACCCGACGATGCTGCTGGGCAACACGACACCTCTGGGCGCCCTGGTGTCT 539
824 ACGTGGGGCGCGCGGGCTGCTGCTAGTGTGTAGCCGTGTGGTGGCGGTGACAAC 883

540 CCAGTGTGTGCCACACTGGCTTCTCCACACGAGACTACATCAATGCCAGATGATCTG 599
884 CCAGTGTGTGCCACAGGCACTCTTCGCTGCCACAGGCACTCCAGCGTAACCGACTGTG 943

600 GTATGACATGACACGACCAAGAAATTTTATGCGGCTTTTTCCTACGCAATGTTCTGAC 659
944 CTATGACCTCAGCGCGCTGCGCTGGCCACCCACTATATGCCCTATGCGATGGCTCTCAC 1003

660 ATTGTCTGGCTTCTT---TTCCCTCTCTGTCATTTTGGTGTGCTTATTCATGATGCTCAG 717
1004 TGTCTATGGCTTCTCTGCTGCGCTTTTGGCTGCGCTGCTGCGCTGCTACTCTCTCTGGCGCTG 1063

718 GAGCCTGTATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCCAGGTC 777
1064 CGCCTGTGTCGCCAGGATGGCCCGGACAGGCTGTGGCCAGGAGCGCGTGGCAAGGC 1123

778 CATCCGACCATCTCTACTGCTGTGTGGCTCTTCAACCTCTGTGTTTGGCTTCCATAT 837
1124 GGCCTGATGCGCGGTGGTGGTGGCTGCTGCTGCTTTCCTCATGACTTCTCTGCTTTTTCAT 1183

838 CACTCGCTCTTCTACCTCACCACCATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGAT 897
1184 CACCAAGACAGGCTACCTGCGAGTGGCTTCGAGCGCGGCGCTCCCTGCACTGATTGGA 1243

898 GGCAGCCAGTGTGGCTCAAGATATGAGGCGCTCTGTGTGAGTGTGAGCAGCTGCTCAA 957
1244 GGCCTTTCAGCGGCTCAAGAGCAGCGCGCTTTCAGAGTGCACAGGCGTCTGGA 1303

958 CCCAGTCTCTGACT 971
1304 CCCATCTCTTCT 1317

RESULT 15
CR598765 1690 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DI069YG10 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR598765
VERSION 1 (bases 1 to 1690)
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1690)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1690
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

/clone="CS0DI069YG10"
/tissue type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT 6"

```

ORIGIN

Query Match	15.1%	Score 163.2	DB.3	Length 1690
Best Local Similarity	51.9%	Pred. No. 4.7e-33		
Matches 443	Conservative 0	Mismatches 403	Indels 8	Gaps 3
Qy	126	CTGCCAGTCTCTCAGAGAAGTACAAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT	185	
Db	464	CTGTGTCTACGGCGAGAACTTCAAGCAACTGCTGCTGCCACCTGTGTATTTCGGCGGTGCT	523	
Qy	186	TATCTTAGGGTGCCTAAATGSCACTGTCTTGTGCACTCTCTTGGCACTCTCTTGGGGCCAAACCAAGCG	245	
Db	524	GGCGGTGGGCTGCGGTGGAACATCTGTGTCTATTACCACAGATCTGCACTGCCCGCGGGC	583	
Qy	246	CTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCACCTGCTTTATGTGCT	305	
Db	584	CTTGACCGCAGCGCGGTGTACACCTTAACCTTGCTCTGGCTGACCTGCTATATGCTG	643	
Qy	306	AT---TGCCTTCTCATCATCATCTACTACTAGATGACAGGTGGGCCCTTCGGGGAGCT	362	
Db	644	CTCCTGCCCTGTCTACTACAACTATGCCAAGGTGATCACTGGCCCTTTTGGCGACTT	703	
Qy	363	GCTCTGAAGCTGTGCACTCTCTGTGTTCTATATCAACTTTTACGGCAGCATCTGCTGCT	422	
Db	704	CGCTGCGGCTGCTCGCTTCTTCTCTATGCCAACTGTCACGGCAGCATCTCTTCCT	763	
Qy	423	GACTGATCTCTGTGCACCAAGTCTCTAGTGTGTGCACCACTGTGTGGCTGCCCTA	482	
Db	764	CACCTGATCAGCTTTCAGCGACTACTTGGGCATCTGCCAACCTGTCACCGCTGCCCGCCGACAA	823	
Qy	483	CC---GGACCGCAGGATGCTCGCTGGGCACCCAGCACCATCTGGGCCCTGTGGTGGTCT	539	
Db	824	ACGTGGGGCGCGCGGCTGCTGGCTAGTGTGTGTAGCCGTGTGGCTGGCCGTGACAAC	883	
Qy	540	CCAGCTCTGCCCACTGCGCTTCTCCCAACGGACTACATCAATGGCCAGATGATCTG	599	
Db	884	CCAGTGCCTGCCACAGCCATCTTCGCTGCCACAGGCATCCAGCGTAACCGCACTGTCTG	943	
Qy	600	GTATGATGACGACCAAGAGATTTTGATCGGCTTTTGGCTACGGCATAGTCTTGAC	659	
Db	944	CTATGACTCAGCCCGCTGCGCTGGGCAACCCACTATATGCGCTATGGCATGGCTCTCAC	1003	
Qy	660	ATTGTCTGGCTTCT--TTCCCTCTTGGTCATTTTGGTGTGCTATTTCACTGATGGTCAG	717	
Db	1004	TGTCATCGGCTTCTGCTGCGCTTGTGCGCTTGTGCGCTGCTACTGTCTCCTGGCGCTG	1063	
Qy	718	GAGCTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCCGACCGCAGTCT	777	
Db	1064	CCGCTGTGCGCGCAGGATGGCCCGGAGAGCTGTGGCCCGAGAGCGGGGTGGCAAGGC	1123	
Qy	778	CATCCGGACCATCTACTGTGTGTGTGGCTCTTCAACCTCTGTGTTTGTGCCCTTCCATAT	837	
Db	1124	GGCCCGCATGGCCGTGGTGGTGGCTGTGCTTTTGCCATCAGCTTCTGCTTTTTCAT	1183	
Qy	838	CATCTGCTCTTTTACCTTCAACATCTCGTCTTCTGCTTTCTCAGGACTTGCAGCTCTTGAT	897	
Db	1184	CACCAAGACAGCTTACCTTGGCAGTGCCTCTGACGCGGGGCTGCCCTGCACTGTATTTGGA	1243	
Qy	898	GGCAGCCAGTGTGGCTTACAAGATATGAGGCGCTCTGTGTAGTGTGAGCAGCTGCCCTCAA	957	
Db	1244	GGCTTTTGACGGGCTTACAAGGCAAGCGGCGCTTTGCCAGTGTCCAAACAGCGTGTGGA	1303	
Qy	958	CCCAGTCTCTTACT	971	
Db	1304	CCCCATCTCTTCT	1317	

Search completed: November 10, 2004, 11:49:19
Job time : 3909 secs

This Page Blank (uspto)